DEFENCE R&D DÉFENSE

A Study of Fast Converging Blind Equalizers

Brian Kozminchuk
Defence Research Establishment Ottawa

Sichun Wang
Telexis Corporation

DEFENCE RESEARCH ESTABLISHMENT OTTAWA

TECHNICAL MEMORANDUM
DREO TM 1999-097
November 1999

National Defence Défense nationale

Canada

101-02396/

DEFENCE R&D DÉFENSE

A Study of Fast Converging Blind Equalizers

Brian Kozminchuk
EW Signal Processing Group
Electronic Support Measures Section

Sichun Wang
Telexis Corporation

DEFENCE RESEARCH ESTABLISHMENT OTTAWA

TECHNICAL MEMORANDUM
DREO TM 1999-097
November 1999

Project
5BB12

ABSTRACT

Two blind identification and equalization algorithms based on second order statistics are investigated. These algorithms have been reported in the literature as having fast convergence, and have been suggested for equalizing signals propagating in time-varying channels. However, two significant difficulties with these algorithms were discovered. First, it was found that they are sensitive to slight inaccuracies in the signal subspace dimension. Standard methods used to estimate the signal subspace dimension provide values which are in the vicinity of the desired one, but which may require adjustment by the user to obtain good blind equalization results. The second difficulty is that the two algorithms enhance noise for low oversampling ratios to such an extent that the transmitted symbol constellation may not be identifiable, even when the SNR is as high as 40 or 50 dB and even when the dimension of the signal subspace is correctly estimated. To overcome this limitation, the oversampling ratio must be increased, which has the desirable effects of decreasing the noise enhancement and improving the convergence rate of the blind equalizer for QAM and M-PSK signals at SNRs as low as 25 to 30 dB. However, this improvement results in a significant increase in computational effort. In contrast, the constant modulus algorithm (CMA) is computationally simpler, but requires longer segments of signal to converge to a set of equalizer coefficients compared to the fast converging algorithms discussed in this report. Thus, for slow-varying channels, the CMA algorithm will likely remain the preferred choice for real time operations. For rapidly-varying channels where the CMA algorithm would fail to equalize the signal, the more computationally intensive fast converging algorithms would have to be used. However, until computers become faster, or the computations are distributed over several processors, the algorithm would have to be relegated to batch or offline applications.

RÉSUMÉ

Deux algorithmes d'identification et d'égalisation aveugles basés sur des statistiques de second ordre sont étudiés. On signale dans la documentation que ces algorithmes convergent rapidement et sont proposés pour égaliser des signaux qui se propagent dans des canaux non-stationnaires. Toutefois, deux difficultés importantes liées à ces algorithmes ont été soulevées. D'abord, on a découvert qu'ils peuvent être sensibles à de légères imprécisions de la dimension sous-spatiale du signal. Les méthodes standard employées pour estimer la dimension sous-spatiale du signal fournissent des valeurs qui sont proches des valeurs vraies, mais qui doivent être ajustées par l'utilisateur pour obtenir de bons résultats d' égalisation à l'aveugle. La deuxième difficulté réside dans le fait que les deux algorithmes accroissent tellement le bruit pour les faibles rapports de suréchantillonnage du signal que la constellation de symboles transmise n'est parfois plus identifiable, même lorsque le SNR est aussi élevé que 40 ou 50 dB, et que la dimension du sous-espace du signal est correctement estimé. Pour surmonter cette limitation, il faut augmenter le rapport de suréchantillonnage, ce qui abaisse le niveau du bruit et permet la convergence rapide de légaliseur aveugle à des SNR de 25 à 30 dB pour les signaux QAM et M-PSK. Cependant, ces améliorations donnent lieu à une augmentation importante de l'effort en calcul. Comparé à l'algorithme à convergence rapide décrit dans ce rapport, l'algorithme module constant (CMA) est plus simple pour ce que est du calcul. Mais en contrepartiel il nécessite des segments de signal plus long pour converger vers un ensemble de coefficients égalisateurs. Donc pour des cannaux à variation lente l'algorithme CMA sera probablement préféré pour des applications en temps reél. Pour des cannaux à variation rapide, où l'algorithme CMA ne réussira pas à égaliser le signal, donc l'algorithme à convergence rapide qui nécessite plus de temps du calcul devra être utilisé. Toutefois, jusqu'à ce que les ordinateurs deviennent plus rapides, ou que les calculs soient partagés entre plusieurs processeurs, cet algorithme devra être utilisé uniquement pour des applications en différé.

EXECUTIVE SUMMARY

A DREO program of research involves the development and assessment of algorithms to blindly equalize and demodulate uncooperative digitally modulated signals of interest. Equalization is potentially useful for correcting distortions in the received signal due to multipath effects and thereby improving the effectiveness of communications ESM systems, such as those being considered for the Army's Land Intelligence and Electronic Warfare Analysis (LIEWA) Project (D6479), in extracting useful information about the signal. The blind equalization problem considered here is complicated by the absence or uncertainty of *a priori* information about the signal of interest, i.e., information such as modulation type, exact carrier frequency, training sequences used for channel equalization, precise symbol timing, basic signal pulse shape, etc.

The work reported on here extends previous work which addressed blind synchronization and equalization issues using the Constant Modulus Algorithm (CMA). It documents performance results of two fast converging blind equalization algorithms based on second order cyclostationary statistics. This study is intended to assess the practical utility of these algorithms relative to the CMA algorithm which is widely used because of its robustness and low computational cost, but which has a slower convergence rate that may cause performance degradation in time-varying channels.

This report documents computer simulations of two fast converging blind identification and equalization algorithms. Identification here refers to the fact that these algorithms first provide an estimate of the channel impulse response (therefore identifying it), followed by an estimate of its reciprocal (used to equalize the signal). The main components of these algorithms are reviewed and simulation results are then discussed. The following simplifying assumptions have been made to minimize the number of variables considered. First, it is assumed that the signal's symbol rate (2400 symbols/sec.) has been estimated prior to the utilization of this algorithm. This allows the signal to be sampled at an integer multiple of the symbol rate, although a time offset (which must be estimated by the equalizer) is assumed to exist between the A/D converter and the optimum sampling instant. Second, it is assumed that the exact carrier frequency of the detected signal is known. Third, the signal intercepted is assumed to be either M-PSK or M-QAM. Finally, a three path multipath channel has been used to assess what the performance capabilities of these equalization algorithms might be.

Two difficulties with these algorithms were discovered. First, it was found that

the algorithms are quite sensitive to slight inaccuracies in the signal subspace dimension. The user must make adjustments to the estimate so that reasonable blind equalization results are obtained. The second difficulty is that the two algorithms enhance noise for low oversampling ratios of the signal to such an extent that the signal constellation is not identifiable even when the SNR is as high as 40 or 50 dB. To overcome this limitation, the oversampling ratio must be increased, which has the desirable effects of decreasing the noise enhancement and improving the convergence rate for QAM and M-PSK signals at SNRs as low as 25 to 30 dB. This improvement in performance, however, is accomplished at the expense of a significant increase in computational effort. In contrast, the constant modulus algorithm (CMA) is computationally simpler, but requires longer segments of signal to converge to a set of equalizer coefficients compared to the fast converging algorithms discussed in this report. Thus, for slow-varying channels, the CMA algorithm will likely remain the preferred choice for real time operations. For rapidly-varying channels where the CMA algorithm would fail to equalize the signal, the more computationally intensive fast converging algorithms would have to be used. However, until computers become faster, or the computations are distributed over several processors, the algorithm would have to be relegated to batch or offline applications.

Kozminchuk, B. W., Wang, S., A Study of Fast Converging Blind Equalizers, Defence Research Establishment Ottawa, DREO TM 1999-097, November, 1999.

SOMMAIRE

Un programme de recherches du CRDO porte sur l'évaluation d'algorithmes visant à égaliser à l'aveugle et à démoduler des signaux modulés numérique-ment, lesquels signaux étant refractaires. L'égalisation peut être utile pour corriger la distorsion dans le signal reçu à la transmission multivoie et, par conséquent, améliorer l'efficacité des systèmes de communications ESM, comme ceux du projet (D6479) de l'Armée, Land Intelligence and Electronic Warfare Analysis (LIEWA), en leur permettant d'extraire l'information utile. Le problème de l'égalisation à l'aveugle dont il est question ici est compliqué par l'absence ou l'inexactitude d'une information *a priori* sur le signal à traiter. C'est-à-dire des renseignements comme le type de modulation, la fréquence exacte de la porteuse, les séquences de formation employées pour l'égalisation des canaux, le minutage précis des symboles, la forme de l'impulsion du signal de base, etc.

Les travaux de présent rapport donnent suite à des travaux antérieurs sur les questions de synchronisation et d'égalisation à l'aveugle utilisant l'algorithmes du module constant (CMA). Ces travaux documentent la performance des algorithmes d'égalisation à l'aveugle à convergence rapide basés sur les statistiques cyclostationnaire de second degré. L'étude vise à évaluer l'utilité pratique de ces algorithmes relativement à l'algorithme CMA qui est couramment utilisé en raison de sa robustesse et de son faible coût en calcul. Toutefois son taux de convergence est plus lent, ce qui pourrait réduire la performance dans les canaux à variation temporelle.

Le présent rapport documente les simulations informatiques de deux algorithmes d'identification et d'égalisation à l'aveugle à convergence rapide. Par identification il doit être compris que ces algorithmes fournissent d'abord une estimation de la réponse en impulsion du canal (l'indentifiant par le fait même), suivie d'une estimation de sa réciproque (utilisée pour égaliser le signal). Les principales composantes de ces algorithmes sont révisées et les résultats de la simulation sont alors discutés. Les hypothèses suivantes ont été formulées pour réduire au minimum le nombre de variables en jeu. Premièrement, on suppose que le débit du signal (2400 symboles/sec.) a été estimé avant l'utilisation de l'algorithme. Ceci permet au signal d'être échantillonné à un taux qui est un multiple entier du taux de symboles quoique l'on suppose qu'il existe un décalage en temps (qui doit être estimé par l'égalisateur) entre le convertisseur A/D et l'instant d'échantillonnage optimal. Deuxièmement, on suppose que la fréquence exacte de la porteuse du signal détecté est connue. Troisièmement, on suppose que le signal intercepté est soit M-PSK ou M-QAM. Enfin, on a utilisé un canal de transmission multiple à trois voies pour évaluer

quelles sont les capacités de ces algorithmes d'égalisation.

Deux difficultés associées à ces algorithmes ont été découvertes. D'abord, on a trouvé que les algorithmes sont plutôt sensibles aux légères inexactitudes dans la dimension sous-spatiale du signal. L'utilisateur doit ajuster l'estimation de manière à obtenir des résultats d'égalisation à l'aveugle raisonnables. La deuxième difficulté est que les deux algorithmes accroissent le bruit pour les faibles rapports de suréchantillonage du signal de telle manière que la constellation du signal n'est pas identifiable lorsque le SNR atteint 40 ou 50 dB. Pour surmonter cette limitation, il faut augmenter le rapport de suréchantillonage, ce qui permettra d'abaisser l'accroissement du bruit et d'atteindre une convergence rapide aux SNR de 25 à 30 dB pour les signaux QAM et M-PSK. Toutefois, cette amélioration de la performance s'accomplit au prix d'une augmentation considérable de l'effort de calcul. Comparé à l'algorithme à convergence rapide décrit dans ce rapport, l'algorithme module constant (CMA) est plus simple pour ce que est du calcul. Mais en contrepartiel il nécessite des segments de signal plus long pour converger vers un ensemble de coefficients égalisateurs. Donc pour des cannaux à variation lente l'algorithme CMA sera probablement préféré pour des applications en temps reél. Pour des cannaux à variation rapide, où l'algorithme CMA ne réussira pas à égaliser le signal, donc l'algorithme à convergence rapide qui nécessite plus de temps du calcul devra être utilisé. Toutefois, jusqu'à ce que les ordinateurs deviennent plus rapides, ou que les calculs soient partagés entre plusieurs processeurs, cet algorithme devra être utilisé uniquement pour des applications en différé.

Kozminchuk, B. W., Wang, S., Étude sur les égalisateurs á l'aveugle á convergence rapide, Le Centre de recherches pour la défense Ottawa, DREO TM 1999-097, Novembre, 1999 (en anglais)

TABLE OF CONTENTS

	Page
ABSTRACT	iii
RÉSUMÉ	iv
EXECUTIVE SUMMARY	v
SOMMAIRE	vii
TABLE OF CONTENTS	ix
LIST OF FIGURES	xi
1.0 INTRODUCTION	1
2.0 THE TXK BLIND EQUALIZATION ALGORITHM	2
2.1 Signal Model	3
2.2 TXK Algorithm Steps	7
2.3 Proof of the TXK Algorithm	10
3.0 THE SS BLIND EQUALIZATION ALGORITHM	18
3.1 Signal Model	18
3.2 SS Algorithm Steps	21
3.3 Proof of the SS Algorithm	22
4.0 EFFECTS OF SIGNAL SUBSPACE DIMENSION ON THE PERFORMANCE OF THE TXK AND SS ALGORITHMS	25
5.0 COMPUTER SIMULATIONS	26
6.0 CONCLUDING REMARKS	43
REFERENCES	44

X

LIST OF FIGURES

		Page
Figure 1:	Examples of raised cosine and square-root raised cosine pulses. . . .	28
Figure 2:	Impulse response of three-path channel.	30
Figure 3:	Examples of signal constellations where the signal has not been equalized.	31
Figure 4:	Equalized signal constellation for TXK algorithm, raised cosine pulse, 16-QAM modulation, block size 300 symbols.	33
Figure 5:	Equalized signal constellation for TXK algorithm, square-root raised cosine pulse, 16-QAM modulation, block size 300 symbols.	34
Figure 6:	Equalized signal constellation for TXK algorithm, 8-PSK, 100 symbols.	35
Figure 7:	Equalized signal constellation for TXK algorithm, square-root raised cosine pulse, 8-PSK modulation, block size 100 symbols.	36
Figure 8:	Equalized signal constellation for SS algorithm, 16-QAM, block size 300 symbols.	37
Figure 9:	Equalized signal constellation for SS algorithm, square-root raised cosine pulse, 16-QAM modulation, block size 300 symbols.	38
Figure 10:	Equalized signal constellation for SS algorithm, 8-PSK, block size 100 symbols.	39
Figure 11:	Equalized signal constellation for SS algorithm, square-root raised cosine pulse, 8-PSK modulation, block size 100 symbols.	40
Figure 12:	Equalized signal constellation for CMA algorithm, raised cosine pulse, equalizer length equal to 12 taps, 16-QAM modulation.	41
Figure 13:	Equalized signal constellation for CMA algorithm, raised cosine pulse, equalizer length equal to 12 taps, 8-PSK modulation.	42

1.0 *INTRODUCTION*

Since the publication of the original papers of Sato [1] and Godard [2], there have appeared many blind identification and equalization algorithms in the literature [3], [6], [7]. Many of these algorithms are based implicitly or explicitly on higher order statistics [2], [8]-[16]. Research efforts on blind equalization algorithms based on higher order statistics have culminated in the so-called universal blind deconvolution methods of O. Shalvi and E. Weinstein [12]. The minimum entropy deconvolution method of Wiggins and Donoho [8] and the constant modulus algorithm (CMA) of Godard [2] (proposed also independently by J. Treichler and B. Agee [9]) have been shown to be important special cases of the universal blind deconvolution techniques in [12]. The CMA algorithm is based on second order and fourth order cumulants of channel outputs and has been extensively studied by researchers. The theoretical analysis provided in [10, 11, 12] explains why the CMA is so successful in equalizing M-PSK signals in a multipath environment. Supported by theoretical analysis, computer simulations and practical implementations, the CMA algorithm is arguably the best adaptive blind equalization algorithm based on higher order statistics developed in the open literature.

In an earlier report [17], a fractionally spaced equalizer updated by the CMA was studied in connection with a blind synchronization algorithm developed in [18]. Preliminary testing demonstrated that the CMA can have difficulty tracking fast varying channels of the type encountered in the HF band. In [18] it was suggested that for such channels, a robust blind equalization algorithm converging faster than the CMA would be useful. The TXK algorithm proposed in [4] and a closely related algorithm proposed in [19] (referred to as the SS algorithm here) seemed to have the desired characteristics of a blind equalization algorithm converging faster than the CMA, since they are based on second order statistics only. For example, in [4] it was shown that the the TXK algorithm could yield accurate channel estimates using only one or two hundred symbols for the simple two-path channel which was considered. It was decided, therefore, to evaluate these algorithms to develop a deeper understanding of their capabilities.

The TXK and SS algorithms are built on certain highly structured (rigid) vector random process models. Theoretically their performance depends on certain regularity conditions which the channel matrix must satisfy and also depends on the correct estimation of the signal subspace dimension. In this report, computer simulations have shown that the standard order estimation methods such as MDL and AIC in [20] provide good estimates of the signal subspace dimension when the oversampling ratio of the signal is

generally significantly higher than the signal's symbol rate. It was found for the three path channel considered that for sampling rates up to four times the symbol rate, both the TXK and SS algorithms have a tendency to enhance noise to such an extent that they fail to resolve the transmitted symbols for SNRs as high as 40 to 50 dB, even when the signal subspace dimension is correctly estimated. To circumvent this problem, it was decided to try a higher sampling rate of the signal to improve the performance, but at the expense of increasing the computational burden. With a proper choice of oversampling rate, convergence rates two to four times faster than the CMA algorithm could be achieved. Nevertheless, because of the computational difficulties associated with the TXK and SS algorithms, they are not practical for on-line applications. However, they could be useful for off-line signal analysis applications.

This report is organized as follows. In Section 2.0, a vector representation of a time invariant baseband communication channel is first established and the TXK algorithm is then formulated and proven. In Section 3.0, another vector representation of a time invariant baseband communications channel is derived for the SS algorithm. The SS algorithm is then formulated and followed by its proof. The effects of signal subspace dimension on the performance of the TXK algorithm and the SS algorithm are briefly discussed in Section 4.0. In Section 5.0 Monte Carlo computer simulation results are presented for both algorithms. Finally, in Section 6.0, some conclusions are drawn.

The notations in this paper follow the original papers [5], [19] closely with some modifications. In particular, if $\mathbf{A}$ is a complex matrix (or vector), then $\mathbf{A}^*$, $\mathbf{A}^T$, $\mathbf{A}^H$ and $\mathbf{A}^\dagger$ denote its complex conjugate, transpose, complex conjugate transpose and pseudoinverse, respectively.

2.0 THE TXK BLIND EQUALIZATION ALGORITHM

This section presents the signal model on which the TXK algorithm is based, the algorithm steps, and a proof of the algorithm.

2.1 Signal Model

When the communication channel is time invariant, the received complex baseband signal $x(\cdot)$ can be expressed as

$$x(t) = r(t) + w(t) \tag{1}$$

where

$$r(t) = \sum_{l=\infty}^{\infty} s_l h(t - lT_s) \tag{2}$$

and the other symbols are defined as follows: s_l represents the information symbols in a signal constellation $\mathcal{S}$; $h(\cdot)$ is the discrete-time composite channel impulse response that includes the pulse shaping filter at the transmitter, the channel, and the receiver filters; T_s is the symbol period; and $w(t)$ is the additive noise.

The following conditions are assumed in the TXK algorithm:

1. The symbol period T_s is known and is an integer multiple of the sampling period Δ;
2. The impulse response $h(\cdot)$ has finite support (finite duration);
3. The symbol sequence $\{s_l\}$ is zero mean and $E(s_k s_l^*) = \delta(k - l)$, where δ is the discrete-time impulse function;
4. $w(\cdot)$ is zero mean, uncorrelated with $\{s_l\}$, and $E(w(t_1)w^*(t_2)) = \sigma^2\delta(t_1 - t_2)$;
5. The channel matrix **H** (to be defined later in this section) is of full column rank.

The vector model upon which the TXK algorithm is based is derived first. Let $t_0 \geq 0$ and $L > 0$ be two non-negative real numbers, where t_0 represents an arbitrary timing offset in the A/D converter and L the length of each observation window (to simplify the analysis t_0 can be set equal to 0) . Let L_h denote the length of the channel impulse response $h(\cdot)$, that is, $h(t) = 0$ if $t < 0$ or $t > L_h$ (note that L and L_h are not necessarily integers, but for simplicity they will be assumed to be integer multiples of the symbol period T_s; also $L < L_h$ typically). Let $[t_0 + kT_s, t_0 + kT_s + L]$ be the boundaries of the k-th observation window. Consider the first window $[t_0, t_0 + L]$. There are only a finite number of symbols s_l contributing to the signal $x(t)$ in the interval $[t_0, t_0 + L]$ because of the finite duration of the channel impulse response $h(\cdot)$. In fact, for any integer l, the function $s_l h(t - lT_s)$

vanishes outside the interval $[lT_s, lT_s + L_h]$ and hence does not contribute to the signal $x(t)$ in $[t_0, t_0 + L]$ if either $lT_s + L_h \leq t_0$ or $lT_s \geq t_0 + L$. This implies that only symbols s_l with index l satisfying the condition $(t_0 - L_h)/T_s \leq l \leq (t_0 + L)/T_s$ contribute to the signal $x(t)$ in the observation window $[t_0, t_0 + L]$. Let $K_0 = (t_0 - L_h)/T_s$ and $K_1 = [(t_0 + L)/T_s]$. The preceding argument shows that in the observation window $[t_0, t_0 + L]$ only the symbols $s_{K_0}, s_{K_0+1}, \cdots, s_{K_1}$ contribute to the signal $x(t)$. If the number of these symbols is denoted by $d = K_1 - K_0 + 1$, then the signal $x(t)$ in the interval $[t_0, t_0 + L]$ can be rewritten as

$$x(t) = \sum_{l=K_0}^{K_1} s_l h(t - lT_s) + w(t) = \sum_{l=K_0}^{K_0+d-1} s_l h(t - lT_s) + w(t). \tag{3}$$

If $x(t)$ in Eq. 3 is sampled in the interval $[t_0, t_0 + L]$ with a sampling period Δ and $L = m\Delta$ where m is the number of samples in the interval $[t_0, t_0 + L]$, then it follows that

$$x(t_0 + i\Delta) = \sum_{l=K_0}^{K_1} s_l h(t_0 + i\Delta - lT_s) = \sum_{l=0}^{d-1} s_{K_0+l} h(t_0 + i\Delta - (K_0 + l)T_s) + w(t_0 + i\Delta), \tag{4}$$

where $1 \leq i \leq m$. Let the $m \times d$ channel matrix $\mathbf{H}(t_0)$ be defined by

$$\mathbf{H}(t_0) = \begin{pmatrix} h(t_0 + \Delta - K_0 T_s) & \cdots & h(t_0 + \Delta - (K_0 + d - 1)T_s) \\ \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots \\ h(t_0 + m\Delta - K_0 T_s) & \cdots & h(t_0 + m\Delta - (K_0 + d - 1)T_s) \end{pmatrix}. \tag{5}$$

The matrix in Eq. 5 can be viewed as follows. Each column is a snapshot of a portion of the impulse response. As one progresses to the right from column to column (i.e., from one symbol period to the next), a different part of the impulse response appears in the column. This is equivalent to the impulse response moving across the observation window, where one sees one symbol period more (or less) of the impulse response. Typically the latter part of the first column will contain zeroes, as will the first part of the last column. Also, for the special case when the duration of the impulse response is an integer number of symbol periods, the number of columns is $d - 1$ as opposed to d, since the first column will be all zeroes. Finally, it will be shown later that, once an estimate of $\mathbf{H}$ is obtained, $h(t)$ can be built up from elements of $\mathbf{H}$.

Let the signal and noise vectors related to the observation window $[t_0, t_0 + L]$ be

defined as

$$\begin{cases} \mathbf{X}(t_0) = [x(t_0+\Delta), \cdots, x(t_0+m\Delta)]^T \\ \mathbf{W}(t_0) = [w(t_0+\Delta), \cdots, w(t_0+m\Delta)]^T \end{cases} \tag{6}$$

and let the vector of symbols associated with the observation window be defined as

$$\mathbf{S}(t_0) = [s_{K_0}, s_{K_0+1}, \cdots, s_{K_0+d-1}]^T. \tag{7}$$

Then the vectors $\mathbf{X}(t_0)$, $\mathbf{W}(t_0)$ and $\mathbf{S}(t_0)$ are related by the equation

$$\mathbf{X}(t_0) = \mathbf{H}(t_0)\mathbf{S}(t_0) + \mathbf{W}(t_0). \tag{8}$$

Similarly, after the k-th symbol period $[t_0 + kT_s, t_0 + kT_s + L]$, if

$$\begin{cases} \mathbf{X}(t_0+kT_s) = [x(t_0+kT_s+\Delta), \cdots, x(t_0+kT_s+m\Delta)]^T \\ \mathbf{W}(t_0+kT_s) = [w(t_0+kT_s+\Delta), \cdots, w(t_0+kT_s+m\Delta)]^T \end{cases} \tag{9}$$

and

$$\mathbf{S}(t_0+kT_s) = [s_{K_0+k}, s_{K_0+k+1}, \cdots, s_{K_0+k+d-1}]^T, \tag{10}$$

we have

$$\begin{aligned} \mathbf{X}(t_0+kT_s) &= \mathbf{H}(t_0+kT_s)\mathbf{S}(t_0+kT_s) + \mathbf{W}(t_0+kT_s) \\ &= \mathbf{H}(t_0)\mathbf{S}(t_0+kT_s) + \mathbf{W}(t_0+kT_s). \end{aligned} \tag{11}$$

Here the fact that $\mathbf{H}(t_0+kT_s) = \mathbf{H}(t_0)$, which is due to the cyclostationary property of the channel [5], has been used. For simplicity $\mathbf{X}(t_0+kT_s)$, $\mathbf{H}(t_0)$, $\mathbf{S}(t_0+kT_s)$, and $\mathbf{W}(t_0+kT_s)$ shall be written as $\mathbf{X}_k$, $\mathbf{H}$, $\mathbf{S}_k$, and $\mathbf{W}_k$ respectively. The vector model in Eq. 11 then becomes

$$\mathbf{X}_k = \mathbf{H}\mathbf{S}_k + \mathbf{W}_k. \tag{12}$$

Equation 12 is the matrix representation of the scalar channel model in Eq. 1 used in the TXK algorithm. With this matrix model, the identification of the channel impulse

response reduces to that of the $m \times d$ channel matrix $\mathbf{H}$, since there is a close correspondence between the channel impulse response function $h(\cdot)$ and the channel matrix $\mathbf{H}$ as explained in [5]. In fact, the channel matrix $\mathbf{H} = \mathbf{H}(t_0)$ is completely determined by the channel impulse response $h(\cdot)$. Conversely, the channel impulse response $h(\cdot)$ can be recovered from the channel matrix $\mathbf{H}$. The exact relationship between the channel impulse response $h(\cdot)$ and the channel matrix $\mathbf{H}$ in Eq. 5 is as follows. Let $\mathbf{H} = \mathbf{H}(t_0)$ be denoted as

$$\mathbf{H} = \begin{pmatrix} h_{11} & h_{12} & \cdots & h_{1(d-1)} & h_{1d} \\ h_{21} & h_{22} & \cdots & h_{2(d-1)} & h_{2d} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ h_{m1} & h_{m2} & \cdots & h_{m(d-1)} & h_{md} \end{pmatrix}. \tag{13}$$

Define a set of column vectors $\mathbf{h}_k$ by

$$\mathbf{h}_k = [h_{1k}, h_{2k}, \cdots, h_{s_r k}]^T , \quad 2 \leq k \leq d \tag{14}$$

where s_r is defined as the number of samples per symbol T_s/Δ (the oversampling factor). This set of column vectors corresponds to all but the first column vector; each vector consisting of the first s_r elements of each vector. To complete the set,

$$\mathbf{h}_1 = [h_{11}, h_{21}, \cdots, h_{s_r 1}, \cdots, h_{m1}]^T. \tag{15}$$

is defined. This vector consists of all elements of the first column. From this set of d column vectors, the desired impulse response vector is formed by concatenating the above column vectors to form the block column vector $\mathbf{h}$ of size $s_r(d-1)+m$, i.e.,

$$\mathbf{h} = [\mathbf{h}_d^T, \mathbf{h}_{d-1}^T, \mathbf{h}_{d-2}^T, \cdots, \mathbf{h}_2^T, \mathbf{h}_1^T]^T. \tag{16}$$

Referring to Eq. 5, it can be verified that the vector $\mathbf{h}$ consists of the samples of the channel impulse response $h(\cdot)$ taken at the equally spaced time instants

$$t_0 + \Delta - (K_0 + d - 1)T_s , \quad t_0 + 2\Delta - (K_0 + d - 1)T_s , \quad \cdots,$$
$$t_0 + \Delta - K_0 T_s , \quad t_0 + 2\Delta - K_0 T_s , \quad \cdots, \; t_0 + m\Delta - K_0 T_s.$$

For example, if the oversampling ratio s_r is half the symbol period (i.e., an oversampling ratio of two samples per symbol period), the channel impulse response is formed by concatenating the top two elements of each column of the matrix in Eq. 13, starting with

the last column. The only exception to this is the first column where all the elements are used.

2.2 TXK Algorithm Steps

The major parameters in the vector representation in Eq. 12 will now be listed followed by the algorithm steps. The notations are very similar to that used in [17].

Main parameters in the TXK algorithm

1. $N = L/T_s$: N is the length of each observation window in symbol periods;
2. $m = L/\Delta$: m is the number of samples in one observation window;
3. $s_r = T_s/\Delta$: s_r is the oversampling ratio (number of samples/symbol) resulting in the relationship $m = s_r N$;
4. $K_0 = [-L_h/T_s]$: $-K_0$ (note that K_0 is negative so that -K_0 is positive) is the duration of the impulse response measured in symbol periods;
5. $d = -(K_0 - 1) + N$: the number of impulse responses (whole or partial) making contributions to the observation window; and
6. σ^2: additive white Gaussian noise power level.

The TXK blind equalization algorithm consists of two steps. First the channel matrix $\mathbf{H}$ is estimated; then the transmitted symbols are recovered. Let $\mathbf{R}_x(0)$ and $\mathbf{R}_x(1)$ be correlation matrices defined by

$$\mathbf{R}_x(0) = E(\mathbf{X}_k \mathbf{X}_k^H) \ , \qquad \mathbf{R}_x(1) = E(\mathbf{X}_k \mathbf{X}_{k+1}^H) \tag{17}$$

where $\mathbf{X}_k^H$ and $\mathbf{X}_{k+1}^H$ (i.e., $\mathbf{X}_k^H$ shifted by one symbol period T_s) are the $m \times 1$ complex conjugate transposes of the vectors $\mathbf{X}_k$ and $\mathbf{X}_{k+1}$ respectively. The TXK algorithm relies only on information contained in the two matrices $\mathbf{R}_x(0)$ and $\mathbf{R}_x(1)$.

Algorithm Steps

1. Select the sampling period Δ with oversampling ratio s_r such that $T_s = s_r\Delta$; select the observation window length L with $L = m\Delta = s_r N\Delta$; then form the $m \times 1$ vectorized observations

$$\mathbf{X}_k = [x(t_0 + \Delta + kT_s), \cdots, x(t_0 + m\Delta + kT_s)]^T$$

where $k = 0, 1, 2, \cdots, m-1$.

2. Select a large positive integer N_1 and estimate $\mathbf{R}_x(0)$ and $\mathbf{R}_x(1)$ from $\mathbf{X}_k$ via time averaging:

$$\begin{cases} \hat{\mathbf{R}}_x(0) = \frac{1}{N_1}\sum_{k=0}^{N_1-1}\mathbf{X}_k\mathbf{X}_k^H \\ \hat{\mathbf{R}}_x(1) = \frac{1}{N_1}\sum_{k=0}^{N_1-1}\mathbf{X}_k\mathbf{X}_{k+1}^H. \end{cases} \tag{18}$$

3. From $\hat{\mathbf{R}}_x(0)$, estimate the noise covariance σ^2 and the dimension d of the signal space using methods described in [20]. More precisely, let $\hat{\mathbf{R}}_x(0)$ have the following singular value decomposition

$$\mathbf{U}_1^H\hat{\mathbf{R}}_x(0)\mathbf{U}_1 = \mathrm{diag}(\beta_1, \cdots, \beta_l, \beta_{l+1}, \cdots, \beta_m) \tag{19}$$

where $\mathbf{U}_1\mathbf{U}_1^H = \mathbf{I}$. Let AIC($k$) and MDL($k$), $0 \le k \le m-1$, be defined by (refer to [20])

$$\mathrm{AIC}(k) = -2\log\left(\frac{\prod_{i=k+1}^{m}\beta_i^{1/(m-k)}}{\frac{1}{m-k}\sum_{i=k+1}^{m}\beta_i}\right)^{(m-k)N_1} + 2k(2m-k) \tag{20}$$

and by

$$\mathrm{MDL}(k) = -\log\left(\frac{\prod_{i=k+1}^{m}\beta_i^{1/(m-k)}}{\frac{1}{m-k}\sum_{i=k+1}^{m}\beta_i}\right)^{(m-k)N_1} + \frac{1}{2}k(2m-k)\log N_1, \tag{21}$$

respectively. Then an estimate $\hat{d}$ of the dimension d of the signal subspace is obtained by minimizing either AIC(k) or MDL(k) and then an estimate $\hat{\sigma}^2$ of the value σ^2 is obtained via the formula

$$\hat{\sigma}^2 = \frac{1}{m-\hat{d}}\sum_{i=\hat{d}+1}^{m}\beta_i. \tag{22}$$

4. Compute the SVD of the estimated noiseless correlation matrix $\mathbf{R}_0$ where

$$\mathbf{R}_0 = \hat{\mathbf{R}}_x(0) - \hat{\sigma}^2\mathbf{I} \tag{23}$$

and form the following: $\mathbf{U}_s$ which consists of the singular vectors associated with the $\hat{d}$ largest singular values; the diagonal matrix $\mathbf{\Sigma}$, whose diagonal consists of the positive square-roots of the $\hat{d}$ largest singular values; and finally the matrix $\mathbf{F} = \mathbf{\Sigma}^{-1}\mathbf{U}_s{}^H$. More precisely, let $\mathbf{R}_0$ have the following singular value decomposition

$$\mathbf{U}^H\mathbf{R}_0\mathbf{U} = \text{diag}(\lambda_1^2, \cdots, \lambda_{\hat{d}}^2, \lambda_{\hat{d}+1}^2, \cdots, \lambda_m^2) \tag{24}$$

where $\mathbf{U}\mathbf{U}^H = \mathbf{I}$ and $\lambda_1 \geq \lambda_2 \geq \cdots \lambda_m$. Let $\mathbf{U} = [\mathbf{u}_1, \cdots, \mathbf{u}_{\hat{d}}, \mathbf{u}_{\hat{d}+1}, \cdots, \mathbf{u}_m]$, where $\mathbf{u}_i$ is the i-th column of $\mathbf{U}$. Then

$$\mathbf{U}_s = [\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_{\hat{d}}] \ , \quad \mathbf{\Sigma} = \text{diag}(\lambda_1, \lambda_2, \cdots, \lambda_{\hat{d}}) \ , \quad \mathbf{F} = \mathbf{\Sigma}^{-1}\mathbf{U}_s{}^H. \tag{25}$$

5. Compute the SVD of $\mathbf{R}$ which is defined as

$$\mathbf{R} = \mathbf{F}\left(\hat{\mathbf{R}}_x(1) - \hat{\sigma}^2\mathbf{J}^{s_r}\right)\mathbf{F}^H, \tag{26}$$

where $\mathbf{J}$ is the $\hat{d} \times \hat{d}$ matrix defined by setting

$$\mathbf{J} = \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix}_{\hat{d}\times\hat{d}} . \tag{27}$$

Also let $\mathbf{y}_{\hat{d}}$ and $\mathbf{z}_{\hat{d}}$ denote the left and right singular vectors corresponding to the smallest singular value.

6. Form an estimate of $\mathbf{H}$ as

$$\hat{\mathbf{H}} = \mathbf{U}_s\mathbf{\Sigma}\mathbf{Q} \tag{28}$$

where

$$\mathbf{Q} = [\mathbf{y}_{\hat{d}}, \mathbf{R}\mathbf{y}_{\hat{d}}, \cdots, \mathbf{R}^{(\hat{d}-1)}\mathbf{y}_{\hat{d}}] \tag{29}$$

or

$$\mathbf{Q} = [(\mathbf{R}^{\dagger})^{(d-1)}\mathbf{z}_d, (\mathbf{R}^{\dagger})^{(d-2)}\mathbf{z}_d, \cdots, \mathbf{z}_d] \tag{30}$$

where $\mathbf{R}^{\dagger}$ is the pseudo inverse of $\mathbf{R}$ from step 5.

7. Recover the information symbol vectors $\mathbf{S}_k$ via

$$\mathbf{S}_k = \hat{\mathbf{H}}^{\dagger}\mathbf{X}_k = \mathbf{Q}^H\mathbf{F}\mathbf{X}_k \tag{31}$$

where $\hat{\mathbf{H}}^{\dagger}$ is the pseudo inverse of $\hat{\mathbf{H}}$. The symbol vectors $\mathbf{S}_k$ can also be recovered via the formula (see Eq. (83) in [17])

$$\mathbf{S}_k = \hat{\mathbf{H}}^H(\hat{\mathbf{H}}\hat{\mathbf{H}}^H + \hat{\sigma}^2\mathbf{I})\mathbf{X}_k. \tag{32}$$

Other approaches can also be used to recover the symbols (see page 347 in [5]).

2.3 Proof of the TXK Algorithm

As indicated in the preceding section, the central problem in the TXK algorithm is the estimation of the channel matrix $\mathbf{H}$ in the vector model from Eq. 12

$$\mathbf{X}_k = \mathbf{H}\mathbf{S}_k + \mathbf{W}_k. \tag{33}$$

It is also assumed that the $m \times d$ channel matrix $\mathbf{H}$ is of full column rank d and the symbol vector $\mathbf{S}_k$ and the noise vector $\mathbf{W}_k$ are wide sense stationary, uncorrelated with each other and zero mean. Hence the sample vector $\mathbf{X}_k$ is also wide sense stationary and zero mean. Moreover, for any integer i, we have

$$\begin{aligned} \mathbf{R}_x(i) &= E[\mathbf{X}_k\mathbf{X}_{i+k}^H] = E\left[(\mathbf{H}\mathbf{S}_k + \mathbf{W}_k)(\mathbf{H}\mathbf{S}_{i+k} + \mathbf{W}_{i+k})^H\right] \\ &= \mathbf{H}E[\mathbf{S}_k\mathbf{S}_{i+k}^H]\mathbf{H}^H + E[\mathbf{W}_k\mathbf{W}_{i+k}^H] \\ &= \mathbf{H}\mathbf{R}_s(i)\mathbf{H}^H + \mathbf{R}_n(i) \end{aligned} \tag{34}$$

where $\mathbf{R}_x(i) = E[\mathbf{X}_k \mathbf{X}_{i+k}^H]$, $\mathbf{R}_s(i) = E[\mathbf{S}_k \mathbf{S}_{i+k}^H]$ and $\mathbf{R}_n(i) = E[\mathbf{W}_k \mathbf{W}_{i+k}^H]$ are the correlation matrices with lag i of the vector processes $\{\mathbf{X}_k\}_{-\infty}^{\infty}$, $\{\mathbf{S}_k\}_{-\infty}^{\infty}$ and $\{\mathbf{W}_k\}_{-\infty}^{\infty}$ respectively. Let $\mathbf{I}_{d\times d}$ be the $d \times d$ identity matrix and define the matrix $\mathbf{J}_{d\times d}$ by

$$\mathbf{J}_{d\times d} = \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix}_{d\times d} . \tag{35}$$

It can be verified that

$$\begin{cases} \mathbf{R}_s(0) = \mathbf{I}_{d\times d} \quad \text{and} \quad \mathbf{R}_n(0) = \sigma^2 \mathbf{I}_{m\times m} \\ \mathbf{R}_s(1) = \mathbf{J}_{d\times d} \quad \text{and} \quad \mathbf{R}_n(1) = \sigma^2 (\mathbf{J}_{m\times m})^{s_r}. \end{cases} \tag{36}$$

The following theorem shows that if the channel matrix $\mathbf{H}$ is of full column rank d, then it is unique up to a complex multiplicative constant:

Theorem 2.0.1 *Suppose there are two matrices $\mathbf{H}_1$, $\mathbf{H}_2$, both of which are of size $m \times d$, of full column rank d and satisfy the linear system in Eq. 33, then there is a constant $c = e^{j\phi}$, where $\phi \in [0, 2\pi)$, such that $\mathbf{H}_2 = c\mathbf{H}_1$.*

Proof. Since $\mathbf{H}_1$ and $\mathbf{H}_2$ both satisfy the linear system in Eq. 33, we have

$$\begin{cases} \mathbf{X}_k = \mathbf{H}_1 \mathbf{S}_k + \mathbf{W}_k \\ \mathbf{X}_k = \mathbf{H}_2 \mathbf{S}_k + \mathbf{W}_k \end{cases} \tag{37}$$

Equations 34 and 36 imply that

$$\begin{cases} \mathbf{R}_x(0) = \mathbf{H}_1 \mathbf{R}_s(0) \mathbf{H}_1^H + \sigma^2 \mathbf{I}_{m\times m} = \mathbf{H}_1 \mathbf{H}_1^H + \sigma^2 \mathbf{I}_{m\times m} \\ \mathbf{R}_x(1) = \mathbf{H}_1 \mathbf{R}_s(1) \mathbf{H}_1^H + \sigma^2 (\mathbf{J}_{m\times m})^{s_r} = \mathbf{H}_1 \mathbf{J}_{d\times d} \mathbf{H}_1^H + \sigma^2 (\mathbf{J}_{m\times m})^{s_r} \\ \mathbf{R}_x(0) = \mathbf{H}_2 \mathbf{R}_s(0) \mathbf{H}_2^H + \sigma^2 \mathbf{I}_{m\times m} = \mathbf{H}_2 \mathbf{H}_2^H + \sigma^2 \mathbf{I}_{m\times m} \\ \mathbf{R}_x(1) = \mathbf{H}_2 \mathbf{R}_s(1) \mathbf{H}_2^H + \sigma^2 (\mathbf{J}_{m\times m})^{s_r} = \mathbf{H}_2 \mathbf{J}_{d\times d} \mathbf{H}_2^H + \sigma^2 (\mathbf{J}_{m\times m})^{s_r} \end{cases} \tag{38}$$

and consequently

$$\begin{cases} \mathbf{H}_1 \mathbf{H}_1^H = \mathbf{H}_2 \mathbf{H}_2^H \\ \\ \mathbf{H}_1 \mathbf{J}_{d\times d} \mathbf{H}_1^H = \mathbf{H}_2 \mathbf{J}_{d\times d} \mathbf{H}_2^H. \end{cases} \tag{39}$$

Since $\mathbf{H}_1$ and $\mathbf{H}_2$ are both of full column rank d, it follows from the first equation in Eq. 39 that there is a unitary matrix $\mathbf{Q}$ of size $d \times d$ such that

$$\mathbf{H}_2 = \mathbf{H}_1\mathbf{Q}. \tag{40}$$

Substituting Eq. 40 into the second equation in Eq. 39 yields

$$\mathbf{H}_1\mathbf{J}_{d\times d}\mathbf{H}_1^H = \mathbf{H}_1\mathbf{Q}\mathbf{J}_{d\times d}\mathbf{Q}^H\mathbf{H}_1^H, \tag{41}$$

or

$$\mathbf{H}_1(\mathbf{J}_{d\times d} - \mathbf{Q}\mathbf{J}_{d\times d}\mathbf{Q}^H)\mathbf{H}_1^H = \mathbf{0}.$$

Since $\mathbf{H}_1$ is of full column rank, we must have

$$\mathbf{J}_{d\times d} - \mathbf{Q}\mathbf{J}_{d\times d}\mathbf{Q}^H = \mathbf{0},$$

or

$$\mathbf{J}_{d\times d}\mathbf{Q} = \mathbf{Q}\mathbf{J}_{d\times d}. \tag{42}$$

The only matrix $\mathbf{Q}$ satisfying Eq. 42 is of the form $\mathbf{Q} = c\mathbf{I}_{d\times d}$, where $c = e^{j\phi}$ for some $\phi \in [0, 2\pi)$ as will now be shown.

Let $\mathbf{Q} = [\mathbf{q}_1, \mathbf{q}_2, \cdots, \mathbf{q}_d]$, where $\mathbf{q}_l$ is the l-th column in the orthogonal matrix $\mathbf{Q}$, $1 \le l \le d$. Then the left side of Eq. 42 can be rewritten as

$$\mathbf{J}_{d\times d}[\mathbf{q}_1, \mathbf{q}_2, \cdots, \mathbf{q}_d] = [\mathbf{J}_{d\times d}\mathbf{q}_1, \mathbf{J}_{d\times d}\mathbf{q}_2, \cdots, \mathbf{J}_{d\times d}\mathbf{q}_d]$$

and the right side rewritten as

$$[\mathbf{q}_1, \mathbf{q}_2, \cdots, \mathbf{q}_d]\begin{pmatrix} 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix}_{d\times d} = [\mathbf{q}_2, \mathbf{q}_3, \cdots, \mathbf{q}_d, 0],$$

hence,

$$\begin{cases} \mathbf{J}_{d\times d}\mathbf{q}_1 &= \mathbf{q}_2 \\ \mathbf{J}_{d\times d}\mathbf{q}_2 &= \mathbf{q}_3 \\ \cdots, &\cdots \quad \cdots \\ \mathbf{J}_{d\times d}\mathbf{q}_{d-1} &= \mathbf{q}_d \\ \mathbf{J}_{d\times d}\mathbf{q}_d &= \mathbf{0} \end{cases} . \tag{43}$$

Let $\mathbf{q}_d = [q_{d,1}, q_{d,2}, \cdots, q_{d,d}]^T$, then the last equation in (43) can be rewritten as

$$\begin{pmatrix} 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix}_{d\times d} \begin{pmatrix} q_{d,1} \\ q_{d,2} \\ q_{d,3} \\ \cdots \\ q_{d,d} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \cdots \\ 0 \end{pmatrix} . \tag{44}$$

Hence $q_{d,1} = q_{d,2} = \cdots = q_{d,d-1} = 0$. Since the vector $\mathbf{q}_d$ has unit norm, $q_{d,d} \neq 0$ and $|q_{d,d}| = 1$. Therefore, there is a number $\phi \in [0, 2\pi)$ such that $q_{d,d} = e^{j\phi}$, that is,

$$\mathbf{q}_d = [0, 0, \cdots, 0, e^{j\phi}]^T.$$

Using this result and then successively solving for $q_d, q_{d-1}, \cdots$ in Eq. 43 yields

$$\begin{cases} \mathbf{q}_{d-1} &= [0, 0, \cdots, e^{j\phi}, 0]^T \\ \mathbf{q}_{d-2} &= [0, 0, \cdots, e^{j\phi}, 0, 0]^T \\ \cdots \quad &\cdots \quad \cdots \quad , \\ \mathbf{q}_2 &= [0, e^{j\phi}, 0, 0, \cdots, 0]^T \\ \mathbf{q}_1 &= [e^{j\phi}, 0, 0, \cdots, 0]^T \end{cases} \tag{45}$$

which proves that $\mathbf{Q} = e^{j\phi}\mathbf{I}_{d\times d}$. Thus, the proof is complete.

Theorem 3.1 shows that as long as the channel matrix $\mathbf{H}$ has full column rank d, it is unique up to a complex multiplicative constant. The next theorem gives a procedure for the construction of the channel matrix $\mathbf{H}$ from the correlation matrices $\mathbf{R}_x(0)$ and $\mathbf{R}_x(1)$ when no additive noise is present. This theorem leads to the TXK algorithm in the general noisy case.

Theorem 2.0.2 *Assume there is no additive noise in the vector random process in Eq. 33,*

that is,

$$\mathbf{X}_k = \mathbf{H}\mathbf{S}_k.$$

Let $\mathbf{R}_x(0) = E(\mathbf{X}_k\mathbf{X}_k^H)$ *and* $\mathbf{R}_x(1) = E(\mathbf{X}_k\mathbf{X}_{k+1}^H)$ *be the correlation matrices of* $\{\mathbf{X}_k\}_{-\infty}^{\infty}$ *(of lags* 0 *and* 1 *symbol period respectively). Let* $\mathbf{R}_x(0)$ *have the singular value decomposition*

$$\mathbf{U}^H\mathbf{R}_x(0)\mathbf{U} = diag(\sigma_1^2, \sigma_2^2, \cdots, \sigma_d^2, 0, 0, \cdots, 0). \tag{46}$$

Let $\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_m]$ *where* $\mathbf{u}_i$, $1 \le i \le m$, *are the columns of* $\mathbf{U}$ *and define the matrices*

$$\begin{cases} \mathbf{U}_s &= [\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_d] \\ \mathbf{\Sigma} &= diag(\sigma_1, \sigma_2, \cdots, \sigma_d) \\ \mathbf{F} &= (\mathbf{\Sigma})^{-1}\mathbf{U}_s{}^H \\ \mathbf{R} &= \mathbf{F}\mathbf{R}_x(1)\mathbf{F}^H. \end{cases} \tag{47}$$

Let $\mathbf{R}$ *have the singular value decomposition*

$$[\mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_d]^H\mathbf{R}[\mathbf{z}_1, \mathbf{z}_2, \cdots, \mathbf{z}_d] = diag(\gamma^2, \gamma_2^2, \cdots, \gamma_d^2) \tag{48}$$

and $\mathbf{Q}$ *be a* $d \times d$ *matrix defined by*

$$\mathbf{Q} = [\mathbf{y}_d, \mathbf{R}\mathbf{y}_d, \mathbf{R}^2\mathbf{y}_d, \mathbf{R}^{d-1}\mathbf{y}_d] \tag{49}$$

then there exists a $\phi \in [0, 2\pi)$ *such that*

$$\mathbf{H} = \mathbf{U}_s\mathbf{\Sigma}\mathbf{Q}e^{j\phi} \tag{50}$$

Proof. Since (refer to Eqs. 34, 36 and 46),

$$\begin{cases} \mathbf{R}_x(0) &= \mathbf{H}\mathbf{H}^H + \sigma^2\mathbf{R}_n(0) = \mathbf{H}\mathbf{H}^H \\ \mathbf{U}^H\mathbf{R}_x(0)\mathbf{U} &= \mathrm{diag}(\sigma_1^2, \sigma_2^2, \cdots, \sigma_d^2, 0, 0, \cdots, 0), \end{cases}$$

we obtain

$$\begin{aligned}
\mathbf{H}\mathbf{H}^H &= \mathbf{U}\mathrm{diag}(\sigma_1^2, \sigma_2^2, \cdots, \sigma_d^2, 0, 0, \cdots, 0)\mathbf{U}^H \\
&= (\mathbf{U}_s, \mathbf{U}_n) \begin{pmatrix} \mathbf{\Sigma}^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} (\mathbf{U}_s, \mathbf{U}_n)^H \\
&= \mathbf{U}_s \mathbf{\Sigma}^2 \mathbf{U}_s{}^H \\
&= \mathbf{U}_s \mathbf{\Sigma} \mathbf{\Sigma}^H \mathbf{U}_s{}^H \\
&= \mathbf{U}_s \mathbf{\Sigma} \left(\mathbf{U}_s \mathbf{\Sigma}\right)^H, \qquad (51)
\end{aligned}$$

where

$$\mathbf{U}_n = [\mathbf{u}_{d+1}, \mathbf{u}_{d+2}, \cdots, \mathbf{u}_m].$$

Hence

$$\mathbf{H}\mathbf{H}^H = \left(\mathbf{U}_s \mathbf{\Sigma}\right) \left(\mathbf{U}_s \mathbf{\Sigma}\right)^H.$$

Letting $\mathbf{G} = \mathbf{U}_s \mathbf{\Sigma}$, the above identity can be simply rewritten as

$$\mathbf{H}\mathbf{H}^H = \mathbf{G}\mathbf{G}^H. \qquad (52)$$

Since $\mathbf{H}$ and $\mathbf{G}$ both have full column rank d, it follows that there is an unitary matrix

$$\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_d] \qquad (53)$$

such that

$$\mathbf{H} = \mathbf{G}\mathbf{V} \qquad (54)$$

or

$$\mathbf{H} = \mathbf{U}_s \mathbf{\Sigma} \mathbf{V}. \qquad (55)$$

It remains to be shown that there is a real number $\phi \in [0, 2\pi)$ such that $\mathbf{V} = \mathbf{Q}e^{j\phi}$. Since $\mathbf{F} = \mathbf{\Sigma}^{-1}\mathbf{U}_s{}^H$ and $\mathbf{R}_x(0) = \mathbf{H}\mathbf{J}_{d\times d}\mathbf{H}^H$ (see Eqs. 47 and 34), then

$$\mathbf{F}\mathbf{H} = \mathbf{\Sigma}^{-1}\mathbf{U}_s{}^H \left(\mathbf{U}_s \mathbf{\Sigma} \mathbf{V}\right) = \mathbf{V} \qquad (56)$$

and

$$\mathbf{R} = \mathbf{F}\mathbf{R}_x(1)\mathbf{F} = \mathbf{F}\mathbf{H}\mathbf{J}_{d\times d}\mathbf{H}^H\mathbf{F}^H = \mathbf{V}\mathbf{J}_{d\times d}\mathbf{V}^H. \qquad (57)$$

Equation 57 can be rewritten as

$$\mathbf{RV} = \mathbf{V}\mathbf{J}_{d\times d} \tag{58}$$

or equivalently

$$[\mathbf{R}\mathbf{v}_1, \cdots, \mathbf{R}\mathbf{v}_d] = [\mathbf{v}_1, \cdots, \mathbf{v}_d] \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix}_{d\times d} . \tag{59}$$

Hence

$$\begin{cases} \mathbf{R}\mathbf{v}_1 &= \mathbf{v}_2 \\ \mathbf{R}\mathbf{v}_2 &= \mathbf{v}_3 \\ \cdots & \cdots \cdots \\ \mathbf{R}\mathbf{v}_{d-1} &= \mathbf{v}_d \\ \mathbf{R}\mathbf{v}_d &= \mathbf{0}. \end{cases} \tag{60}$$

It follows that $\mathbf{v}_2 = \mathbf{R}\mathbf{v}_1, \cdots, \mathbf{v}_d = \mathbf{R}^{d-1}\mathbf{v}_1$. Clearly, once the vector $\mathbf{v}_1$ is determined, the whole matrix $\mathbf{V}$ will be determined. To do that, the matrix $\mathbf{R}\mathbf{R}^H$ must first be computed. In fact, since $\mathbf{R} = \mathbf{V}\mathbf{J}_{d\times d}\mathbf{V}^H$ (refer to Eq. 57) , then

$$\begin{aligned} \mathbf{R}\mathbf{R}^H &= \mathbf{V}\mathbf{J}_{d\times d}\mathbf{V}^H\mathbf{V}\mathbf{J}_{d\times d}^H\mathbf{V}^H \\ &= \mathbf{V}\mathbf{J}_{d\times d}\mathbf{J}_{d\times d}^H\mathbf{H}^H \\ &= \mathbf{V}\mathrm{diag}(0,1,1,1,\cdots,1)\mathbf{V}^H. \end{aligned} \tag{61}$$

Moreover, from Eq. 48

$$\mathbf{R} = [\mathbf{y}_1, \cdots, \mathbf{y}_d]\mathrm{diag}(\gamma_1^2, \cdots, \gamma_d^2)[\mathbf{z}_1, \cdots, \mathbf{z}_d]^H \tag{62}$$

and therefore

$$\begin{aligned} \mathbf{R}\mathbf{R}^H &= [\mathbf{y}_1, \cdots, \mathbf{y}_d]\mathrm{diag}(\gamma_1^2, \cdots, \gamma_d^2)[\mathbf{z}_1, \cdots, \mathbf{z}_d]^H \\ &\quad \times [\mathbf{z}_1, \cdots, \mathbf{z}_d]\mathrm{diag}(\gamma_1^2, \cdots, \gamma_d^2)[\mathbf{y}_1, \cdots, \mathbf{y}_d]^H \\ &= [\mathbf{y}_1, \cdots, \mathbf{y}_d]\mathrm{diag}(\gamma_1^4, \cdots, \gamma_d^4)[\mathbf{y}_1, \cdots, \mathbf{y}_d]^H \end{aligned} \tag{63}$$

where the characteristic roots have been written in decreasing order. Equations 61 and 63 show that 0 is a singular value of $\mathbf{R}\mathbf{R}^H$ with multiplicity 1 and $\mathbf{y}_d$ and $\mathbf{v}_1$ are singular vectors of $\mathbf{R}\mathbf{R}^H$ associated with the singular value 0. It follows that the two vectors $\mathbf{y}_d$ and $\mathbf{v}_1$ are linearly related, that is, there is a complex number β such that $\mathbf{v}_1 = \beta\mathbf{y}_d$. Since $\mathbf{y}_d$ and $\mathbf{v}_1$ both have unit norm, there is a real number $\phi \in [0, 2\pi)$ with $\beta = e^{j\phi}$. We have finally proved that

$$\mathbf{V} = [\mathbf{y}_d, \mathbf{R}\mathbf{y}_d, \cdots, \mathbf{R}^{d-1}\mathbf{y}_d]e^{j\phi} = \mathbf{Q}e^{j\phi}.$$

This completes the proof.

When there is additive noise, the correlation matrices $\mathbf{R}_x(0)$ and $\mathbf{R}_x(1)$ of the observed sample vector $\{\mathbf{X}_k\}$ become

$$\begin{cases} \mathbf{R}_x(0) &= \mathbf{H}\mathbf{H}^H + \sigma^2\mathbf{I}_{m\times m} \\ \mathbf{R}_x(1) &= \mathbf{H}\mathbf{J}_{d\times d}\mathbf{H}^H + \sigma^2(\mathbf{J}_{m\times m})^{s_r}. \end{cases} \tag{64}$$

Let $\mathbf{R}_0 = \mathbf{H}\mathbf{H}^H$ and $\mathbf{R}_1 = \mathbf{H}\mathbf{J}_{d\times d}\mathbf{H}^H$. To reconstruct $\mathbf{H}$ as specified by Theorem 2.0.2, the matrices $\mathbf{R}_0$ and $\mathbf{R}_1$, which correspond to the correlation matrices without noise, must be computed . Let the singular value decomposition of $\mathbf{R}_0$ be

$$\mathbf{R}_0 = \mathbf{U}\mathrm{diag}(\lambda_1^2, \lambda_2^2, \cdots, \lambda_d^2, 0, 0, \cdots, 0)\mathbf{U}^H, \tag{65}$$

then $\mathbf{R}_x(0)$ has the singular value decomposition

$$\mathbf{R}_x = \mathbf{U}\mathrm{diag}(\lambda_1^2 + \sigma^2, \lambda_2^2 + \sigma^2, \cdots, \lambda_d^2 + \sigma^2, \sigma^2, \cdots, \sigma^2)\mathbf{U}^H \tag{66}$$

In theory, from the SVD decomposition (Eq. 66) the signal subspace dimension d can be estimated using a threshold test or the information theoretic criteria (MDL or AIC) as explained in [20], and the noise variance σ^2 can be estimated as the mean of the $m - d$ smallest eigenvalues. Once d and σ^2 are obtained,

$$\begin{cases} \mathbf{R}_0 &= \mathbf{H}\mathbf{H}^H = \mathbf{R}_x(0) - \sigma^2\mathbf{I}_{m\times m} \\ \\ \mathbf{R}_1 &= \mathbf{H}\mathbf{J}_{d\times d}\mathbf{H}^H = \mathbf{R}_x(1) - \sigma^2(\mathbf{J}_{m\times m})^{s_r}. \end{cases} \tag{67}$$

Now the procedure specified in Theorem 2.02 (refer to Eqs. 46 to 50) can be used to obtain an estimate $\hat{\mathbf{H}}$ of the channel matrix $\mathbf{H}$. This completes the derivation of the TXK

algorithm.

3.0 THE SS BLIND EQUALIZATION ALGORITHM

This section describes the signal model of the SS algorithm, the SS algorithm steps, and the proof the algorithm.

3.1 Signal Model

The SS algorithm is similar to the TXK algorithm in that it uses only second order statistics and is also based on a vector model representation of the communications channel defined in Eq. 1. In fact, the vector model used in the SS algorithm can be derived using the vector model in Eq. 12. Here, it is assumed that the conditions which were used in the TXK algorithm also hold. Let L_h be the length of the channel impulse response $h(\cdot)$, the observation window L be equal to the symbol period T_s, $K_0 = (t_0 - L_h)/T_s$ and $K_1 = [(t_0 + L)/T_s]$. Let $d = K_1 - K_0 + 1$, which is the number of impulse responses wholly or partially contained in the observation window. Let $x_k^{(i)} = x(t_0 + i\Delta + kT_s)$, $1 \leq i \leq s_r$, where s_r is the oversampling ratio (i.e., the number of samples per symbol) and $x_k^{(i)}$ is the ith sample from the kth symbol. Then the vector model in Eq. 12 for this special case becomes

$$\begin{pmatrix} x_k^{(1)} \\ x_k^{(2)} \\ x_k^{(3)} \\ \cdots \\ x_k^{(s_r)} \end{pmatrix} = \begin{pmatrix} h(t_0+\Delta-K_0T_s) & \cdots & h(t_0+\Delta-(K_0+d-1)T_s) \\ \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots \\ h(t_0+s_r\Delta-K_0T_s) & \cdots & h(t_0+s_r\Delta-(K_0+d-1)T_s) \end{pmatrix} \begin{pmatrix} s_{K_0+k} \\ s_{K_0+k+1} \\ s_{K_0+k+2} \\ \cdots \\ s_{K_0+k+d-1} \end{pmatrix}$$

$$+ \begin{pmatrix} w(t_0+\Delta+kT_s) \\ w(t_0+2\Delta+kT_s) \\ w(t_0+3\Delta+kT_s) \\ \cdots \\ w(t_0+s_r\Delta+kT_s). \end{pmatrix} \tag{68}$$

As with the TXK algorithm, the columns of the impulse response matrix above consist of

segments of the impulse response as seen in the observation window. The number of impulse response segments is equal to d, i.e., the number of symbols affecting the observation window. A sample $x_k^{(i)}$ taken from the observation window $L = T_s$ consists of a sample of noise plus a weighted sum of impulse responses partially within L. Equation 68 converts the one-channel model of Eq. 1 into s_r subchannels. The channel impulse response of the i-th subchannel is given by the ith row, i.e.,

$$\begin{aligned} \mathbf{h}^{(i)} &= [h_0^{(i)}, h_1^{(i)}, \cdots, h_{d-1}^{(i)}]^T \\ &= [h(t_0 + i\Delta - K_0 T_s) \cdots h(t_0 + i\Delta - (K_0 + d - 1)T_s). \end{aligned} \tag{69}$$

To conform to the notational conventions in the original paper [19], we let $M = d - 1$.

Now let the observation window be expanded to $L = NT_s$ where $N \geq 1$ represents the length of the observation window in symbol periods T_s. Taking the i^{th} sample from each of the N symbol periods, or from each sub-channel, gives the vector

$$\mathbf{X}_k^{(i)} = [x_k^{(i)}, x_{k-1}^{(i)}, \cdots, x_{k-N+1}^{(i)}]^T \quad , \quad 0 \leq i \leq s_r - 1. \tag{70}$$

This gives the ith sub-channel output. Let the $N \times (M + N)$ sub-channel matrix $\mathbf{H}_N^{(i)}$ be defined by

$$\mathbf{H}_N^{(i)} = \begin{pmatrix} h_0^{(i)} & \cdots & h_M^{(i)} & 0 & \cdots & \cdots & 0 \\ 0 & h_0^{(i)} & \cdots & h_M^{(i)} & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & 0 & h_0^{(i)} & \cdots & h_M^{(i)} \end{pmatrix}. \tag{71}$$

It can be verified that the sub-channel matrix equation is

$$\mathbf{X}_k^{(i)} = \mathbf{H}_N^{(i)} \mathbf{S}_k + \mathbf{W}_k^{(i)} \tag{72}$$

where

$$\begin{cases} \mathbf{S}_k = [s_k, \cdots, s_{k-N-M+1}]^T \\ \mathbf{W}_k^{(i)} = [w(i\Delta + kT_s), \cdots, w(i\Delta + (k - N + 1)T_s)]^T. \end{cases} \tag{73}$$

Stacking the s_r sub-channel equation systems in Eq. 73 together, the vector representation

of the channel similar to Eq. 1 is obtained, i.e.,

$$\mathbf{X}_k = \mathbf{H}_N \mathbf{S}_k + \mathbf{W}_k \tag{74}$$

where

$$\begin{cases} \mathbf{X}_k = [(\mathbf{X}_k^{(0)})^T \cdots (\mathbf{X}_k^{(s_r-1)})^T]^T \\ \mathbf{H}_N = [(\mathbf{H}_N^{(0)})^T \cdots (\mathbf{H}_N^{(s_r-1)})^T]^T \\ \mathbf{W}_k = [(\mathbf{W}_k^{(0)})^T \cdots (\mathbf{W}_k^{(s_r-1)})^T]^T \end{cases} \tag{75}$$

The vectors $\mathbf{X}_k$ and $\mathbf{W}_k$ have dimensions $s_r N \times 1$ where s_r is the oversampling ratio and is an integer multiple of the symbol rate $1/T_s$, and $N = L/T_s$.

The vector model in Eq. 74 is different from the vector model in Eq. 12 in that the matrix $\mathbf{H}_N$ in Eq. 74 has more structure than the channel matrix $\mathbf{H}$ in Eq. 12, i.e., the channel here is subdivided into s_r sub-channels. The unique structure of $\mathbf{H}_N$ is used in an essential way in the derivation of the SS algorithm (see Lemma 1 in [19]). To formulate the SS algorithm, additional notations are needed. Let $\mathbf{h}$ be the $s_r(M+1) \times 1$ vector of sub-channel coefficients (recall that $M+1 = d$, and d refers to the number of partial or entire impulse responses contained in the observation window) and is defined as

$$\mathbf{h} = [(\mathbf{h}^{(0)})^T, \cdots, (\mathbf{h}^{(s_r-1)})^T]^T. \tag{76}$$

Let

$$\mathbf{R}_x = E(\mathbf{X}_k \mathbf{X}_k^H) \tag{77}$$

be the autocorrelation matrix of the random vector $\mathbf{X}_k$ (of dimension $s_r N \times 1$). For any given $s_r N \times 1$ vector $\mathbf{v} = [v_1, v_2, \cdots, v_{s_r N}]$ (the elements of $\mathbf{v}$ will be defined later), construct a sequence of $(M+1) \times (M+N)$ Toeplitz matrices $\mathcal{V}_{(M+1)}^{(l)}$, $1 \le l \le s_r$, as follows (see Eqs. 18 and 19 in [19])

$$\mathcal{V}_{(M+1)}^{(l)} = \begin{pmatrix} v_{(l-1)N+1} & v_{(l-1)N+2} & \cdot & \cdot & v_{lN} & 0 & \cdot & \cdot & 0 \\ 0 & v_{(l-1)N+1} & v_{(l-1)N+2} & \cdot & \cdot & v_{lN} & 0 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & 0 & v_{(l-1)N+1} & v_{(l-1)N+2} & \cdot & \cdot & v_{lN} \end{pmatrix} \tag{78}$$

Specifically, the vector $\mathbf{v}$ is divided into s_r sub-vectors of length N elements each. The

l^{th} sub-vector is inserted into $M+1$ successive rows as shown in Eq. 78 giving rise to the $(M+1)\times(M+N)$ matrix. The s_r matrices $\mathcal{V}_{(M+1)}^{(l)}$ are stacked to obtain the matrix

$$\mathcal{V}_{(M+1)} = [(\mathcal{V}_{(M+1)}^{(1)})^T, \cdots, (\mathcal{V}_{(M+1)}^{(s_r)})^T]^T. \tag{79}$$

The matrix $\mathcal{V}_{(M+1)}$ has dimensions $s_r(M+1)\times(M+N)$ and is called the filtering matrix associated with the $s_rN\times 1$ vector $\mathbf{v}$.

Critical to the derivation of the SS algorithm is the identity (Lemma 1 in [19])

$$\mathbf{v}^T\mathbf{H}_N = \mathbf{h}^T\mathcal{V}_{(M+1)}. \tag{80}$$

3.2 SS Algorithm Steps

1. Select the sampling period $\Delta > 0$ with oversampling ratio s_r such that $T_s = s_r\Delta$. Select a positive integer $N \geq 2$ (the number of symbol periods in an observation window) and form the vectorized observations $\mathbf{X}_k$ according to definitions in Eqs. 70 and 75, where $k = 0, 1, 2, \cdots$.

2. Select a large positive integer N_1 representing the number of symbols in a block to be processed, and estimate $\mathbf{R}_x$ (see Eq. 77)) from $\mathbf{X}_k$ via time averaging, i.e.,

$$\hat{\mathbf{R}}_x = \frac{1}{N_1}\sum_{k=0}^{N_1-1}\mathbf{X}_k\mathbf{X}_k^H. \tag{81}$$

3. Compute the SVD of the estimated correlation matrix $\mathbf{R}_x$, estimate the noise covariance σ^2 and the dimension $M+N$ (i.e., $d-1+N$) of the signal space using methods described in [20] (refer to Eqs. 19-21) and then form the signal subspace matrix $\mathbf{U}_s$ and noise subspace matrix $\mathbf{U}_n$ as follows. First let $\mathbf{R}_x$ have the singular value decomposition

$$\mathbf{U}^H\mathbf{R}_x\mathbf{U} = \mathrm{diag}(\lambda_1^2, \cdots, \lambda_{M+N}^2, \lambda_{M+N+1}^2, \cdots, \lambda_{s_rN}^2). \tag{82}$$

Form the MDL(k) or AIC(k) sequence specified by Eqs. 19 to 21 and estimate the dimension $M+N$ of the signal subspace dimension by minimizing MDL(k) or AIC(k) (M becomes s_rN in this case, i.e., the number of samples in the observation window). Let $\mathbf{U} = [\mathbf{u}_1, \cdots, \mathbf{u}_{M+N}, \mathbf{u}_{M+N+1}, \cdots, \mathbf{u}_{s_rN}]$, where $\mathbf{u}_i$ is the i-th column

of $\mathbf{U}$. Then

$$\mathbf{U}_s = [\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_{M+N}] \tag{83}$$

and

$$\mathbf{U}_n = [\mathbf{u}_{M+N+1}, \mathbf{u}_{M+N+2}, \cdots, \mathbf{u}_{s_r N}]. \tag{84}$$

4. Compute the filtering matrices $\mathbf{G}_i$ (which replaces the general form represented by Eq. 79), associated with the $s_r N \times 1$ noise vectors $\mathbf{u}_i$ (which replaces the general form $\mathbf{v}$ in Eq. 78), where $M + N + 1 \leq i \leq s_r N$ from Eq. 84, and, as will be shown in the next section, form the quadratic $q(\mathbf{h}) = \mathbf{h}^T \mathcal{Q} \mathbf{h}^*$, where $\mathbf{h}^*$ is the complex conjugate of the impulse response vector $\mathbf{h}$ and

$$\mathcal{Q} = \sum_{i=M+N+1}^{s_r N} \mathbf{G}_i^* \mathbf{G}_i^T. \tag{85}$$

5. As will be shown in the next section, minimize $q(\mathbf{h})$ subject to the constraint $|\mathbf{h}| = 1$ and obtain an estimate of the channel coefficient vector $\mathbf{h}$.

6. Construct the matrix $\mathbf{E}$ from

$$\mathbf{E} = \mathbf{H}_N^H \mathbf{U}_s \mathrm{diag}((\lambda_1^2 - \hat{\sigma}^2)^{-1}, \cdots, (\lambda_{M+N}^2 - \hat{\sigma}^2)^{-1}) \mathbf{U}_s{}^H. \tag{86}$$

and then recover the symbol vectors $\mathbf{S}_k$ via

$$\mathbf{S}_k = \mathbf{E}\mathbf{X}_k. \tag{87}$$

3.3 Proof of the SS Algorithm

The SS algorithm is based on the orthogonality relation between the so-called signal subspace and noise subspace and the unique structure of the channel matrix $\mathbf{H}_N$ in the model described by Eq. 74. The main ideas involved in the SS algorithm will be explained, but some details are omitted due to space considerations. It is assumed that the channel matrix $\mathbf{H}_N$ is of full column rank $M + N$. By assumption, the symbol vector $\mathbf{S}_k$ and the noise vector $\mathbf{W}_k$ are wide sense stationary, uncorrelated with each other and zero mean. Hence the channel output sample vector $\mathbf{X}_k$ is wide sense stationary and

its autocorrelation matrix $\mathbf{R}_x = E[\mathbf{X}_k\mathbf{X}_k]^H$ with dimensions $s_rN \times s_rN$ can be easily computed from

$$\begin{aligned}\mathbf{R}_x &= E[\mathbf{X}_k\mathbf{X}_k^H] = E\left[(\mathbf{H}_N\mathbf{S}_k + \mathbf{W}_k)(\mathbf{H}_N\mathbf{S}_k + \mathbf{W}_k)^H\right] \\ &= \mathbf{H}_N E[\mathbf{S}_k\mathbf{S}_k^H]\mathbf{H}_N^H + E[\mathbf{W}_k\mathbf{W}_k^H] \\ &= \mathbf{H}_N\mathbf{R}_s\mathbf{H}_N^H + \mathbf{R}_n \end{aligned} \tag{88}$$

where $\mathbf{R}_s = E[\mathbf{S}_k\mathbf{S}_k^H]$ and $\mathbf{R}_n = E[\mathbf{W}_k\mathbf{W}_k^H]$. In this report, as in the original paper [19], it is assumed that $\mathbf{R}_s$ is of full rank $M+N$ and $\mathbf{R}_n = E[\mathbf{W}_k\mathbf{W}_k] = \sigma^2\mathbf{I}_{s_rN\times s_rN}$, where σ^2 is the noise power. The SS algorithm relies only on information contained in the autocorrelation matrix $\mathbf{R}_x$ to estimate the channel coefficient vector $\mathbf{h}$ (defined in Eq. 76).

Since $\mathbf{H}_N$ and $\mathbf{R}_s$ are of full column rank $M+N$, the rank of the matrix $\mathbf{H}_N\mathbf{R}_s\mathbf{H}_N^H$ is also $M+N$ and it has a singular value decomposition of the form

$$\mathbf{U}^H\mathbf{H}_N\mathbf{R}_s\mathbf{H}_N^H\mathbf{U} = \mathrm{diag}(\lambda_1^2, \lambda_2^2, \cdots, \lambda_{M+N}^2, 0, 0, \cdots, 0). \tag{89}$$

The autocorrelation matrix $\mathbf{R}_x$ has the singular value decomposition

$$\mathbf{U}^H\mathbf{R}_x\mathbf{U} = \mathrm{diag}(\lambda_1^2+\sigma^2, \lambda_2^2+\sigma^2, \cdots, \lambda_{M+N}^2+\sigma^2, \sigma^2, \sigma^2, \cdots, \sigma^2). \tag{90}$$

Let

$$\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_{M+N}, \mathbf{u}_{M+N+1}, \mathbf{u}_{M+N+2}, \cdots, \mathbf{u}_{s_rN}] \tag{91}$$

where $\mathbf{u}_l$ is the l-th column of the orthogonal matrix $\mathbf{U}$, $1 \le l \le s_rN$. Let

$$\begin{cases} \mathbf{U}_s &= [\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_{M+N}] \\ \mathbf{U}_n &= [\mathbf{u}_{M+N+1}, \mathbf{u}_{M+N+2}, \cdots, \mathbf{u}_{s_rN}] \end{cases} \tag{92}$$

The subspace spanned by the vectors $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_{M+N}$, denoted by $\mathbf{V}_s$, is called the signal subspace and the subspace spanned by the vectors $\mathbf{u}_{M+N+1}, \mathbf{u}_{M+N+2}, \cdots, \mathbf{u}_{s_rN}$, denoted by $\mathbf{V}_n$, is called the noise subspace. The two subspaces $\mathbf{V}_s$ and $\mathbf{V}_n$ are orthogonal complements to each other. The following result shows that the signal subspace $\mathbf{V}_s$ coincides with the subspace spanned by the columns of the channel matrix $\mathbf{H}_N$.

Theorem 3.0.1 *If we denote the subspace spanned by the columns of the channel matrix*

$\mathbf{H}_N$ *by* $\mathbf{V}_c$, *then* $\mathbf{V}_s = \mathbf{V}_c$.

Proof. It follows from Eq. 90 that

$$\begin{cases} \mathbf{H}_N\mathbf{R}_s\mathbf{H}_N^H\mathbf{u}_1 &= \lambda_1^2\mathbf{u}_1 \\ \mathbf{H}_N\mathbf{R}_s\mathbf{H}_N^H\mathbf{u}_2 &= \lambda_2^2\mathbf{u}_2 \\ \cdots \quad \cdots & \cdots \quad \cdots \quad \cdots \\ \mathbf{H}_N\mathbf{R}_s\mathbf{H}_N^H\mathbf{u}_{M+N} &= \lambda_{M+N}^2\mathbf{u}_{M+N} \end{cases} \tag{93}$$

so that Eq. 93 can be rewritten as

$$\begin{cases} \mathbf{u}_1 &= \frac{1}{\lambda_1^2}\mathbf{H}_N\mathbf{R}_s\mathbf{H}_N^H\mathbf{u}_1 \\ \mathbf{u}_2 &= \frac{1}{\lambda_2^2}\mathbf{H}_N\mathbf{R}_s\mathbf{H}_N^H\mathbf{u}_2 \\ \cdots \quad \cdots & \cdots \quad \cdots \quad \cdots \\ \mathbf{u}_{M+N} &= \frac{1}{\lambda_{M+N}^2}\mathbf{H}_N\mathbf{R}_s\mathbf{H}_N^H\mathbf{u}_{M+N}. \end{cases} \tag{94}$$

Equation 94 shows that the vectors $\mathbf{u}_1, \cdots, \mathbf{u}_{M+N}$, are linear combinations of the column vectors of $\mathbf{H}_N$ since the terms $\mathbf{R}_s\mathbf{H}_N^H\mathbf{u}_i$ are column vectors. Thus the subspace $\mathbf{V}_s$ lies in the subspace $\mathbf{V}_c$. On the other hand, since the column rank of $\mathbf{H}_N$ is $M+N$, the dimension of $\mathbf{V}_c$ is $M+N$. The subspace $\mathbf{V}_s$ is spanned by $M+N$ orthogonal vectors $\mathbf{u}_1, \cdots, \mathbf{u}_{M+N}$ and its dimension is also $M+N$. Clearly the signal subspace spanned by $\mathbf{V}_s$ is the same as that spanned by $\mathbf{V}_c$ since they have the same dimension $M+N$. This completes the proof.

A basic observation made in [19] is that the channel matrix $\mathbf{H}_N$ is completely determined by the subspace $\mathbf{V}_c$ and, therefore $\mathbf{V}_s$. More specifically, suppose that there are two channel matrices $\mathbf{H}_N^1$ and $\mathbf{H}_N^2$ satisfying Eq. 74, such that

$$\begin{cases} \mathbf{X}_k &= \mathbf{H}_N^1\mathbf{S}_k + \mathbf{W}_k \\ \mathbf{X}_k &= \mathbf{H}_N^2\mathbf{S}_k + \mathbf{W}_k \end{cases} \tag{95}$$

where the channel matrices $\mathbf{H}_N^1$ and $\mathbf{H}_N^2$ have the same dimensions and are constructed according to Eqs. 69 and 71. Let the subspaces $\mathbf{V}_c{}^1$ and $\mathbf{V}_c{}^2$ be generated by the columns of $\mathbf{H}_N^1$ and $\mathbf{H}_N^2$ respectively. Then, it is shown in [19] that $\mathbf{H}_N^2 = \beta\mathbf{H}_N^1$, i.e., the two matrices are related to one another through the complex constant β. This result is similar to the one obtained for the TXK algorithm, although via a different approach. Thus, once the signal subspace $\mathbf{V}_s$ is determined from Eqs. 90 to 92, one can use the

noise subspace $\mathbf{V}_n$ to obtain a solution for the channel impulse response, since the noise vectors $\mathbf{u}_{M+N}, \cdots, \mathbf{u}_{s_rN-1}$ are all orthogonal to the columns of the channel matrix $\mathbf{H}_N$. Thus, the following simple relationship holds

$$\mathbf{U}_n{}^H \mathbf{H}_N = \mathbf{0}. \tag{96}$$

It follows that the channel matrix $\mathbf{H}_N$ can be obtained by solving Eq. 96 using a constrained least squares approach since in practice only estimates of the noise vectors $\mathbf{u}_{M+N+1}, \cdots, \mathbf{u}_{s_rN}$ are available. A quadratic form $q(\mathbf{h})$ is formed and then minimized to obtain an estimate of the channel coefficient vector $\mathbf{h}$. More specifically, let the squared error be represented by

$$q(\mathbf{h}) = \sum_{l=M+N+1}^{s_rN} |\mathbf{u}_l^H \mathbf{H}_N|^2. \tag{97}$$

The SS algorithm estimates the vector $\mathbf{h}$ (or the channel matrix $\mathbf{H}_N$) by minimizing $q(\mathbf{h})$.

The quadratic form defined by Eq. 97 is not expressed in terms of $\mathbf{h}$ directly. An important simplification is obtained by using Eq. 80. Let the filtering matrix associated with the $s_rN \times 1$ noise vector $\mathbf{u}_l$ be denoted by $\mathbf{G}_l$, where $M+N+1 \leq l \leq s_rN$ (defined in Eqs. 78 and 79). Using Eq. 80, the quadratic form $q(\mathbf{h})$ can be simplified, i.e.,

$$\begin{aligned} q(\mathbf{h}) &= \sum_{l=M+N+1}^{s_rN} |\mathbf{u}_l^H \mathbf{H}_N|^2 = \sum_{l=M+N+1}^{s_rN} |(\mathbf{u}_l^*)^T \mathbf{H}_N|^2 \\ &= \sum_{l=M+N+1}^{s_rN} |\mathbf{h}^T \mathbf{G}_l^*|^2 = \sum_{l=M+N+1}^{s_rN} \left(\mathbf{h}^T \mathbf{G}_l^*\right) \left(\mathbf{h}^T \mathbf{G}_l^*\right)^H \\ &= \sum_{l=M+N+1}^{s_rN} \mathbf{h}^T \mathbf{G}_l^* \mathbf{G}_l^T \mathbf{h}^* = \mathbf{h}^T \sum_{l=M+N+1}^{s_rN} \mathbf{G}_l^* \mathbf{G}_l^T \mathbf{h}^* \\ &= \mathbf{h}^T \mathcal{Q} \mathbf{h}^* \end{aligned} \tag{98}$$

This completes the proof.

4.0 EFFECTS OF SIGNAL SUBSPACE DIMENSION ON THE PERFORMANCE OF THE TXK AND SS ALGORITHMS

The TXK and SS algorithms are algebraic in nature. Many linear algebraic operations are involved in the equalization process in both algorithms. Since singular value

decomposition and the computation of the pseudo inverse can lead to unreliable results if the channel matrices are ill-conditioned, difficulties with the TXK and SS algorithms can arise. The TXK algorithm relies on the correct estimation of the signal subspace dimension $d = -K_0 + 1 + N$. If the steps in the algorithm are examined, the third step, which determines the signal subspace dimension d, sets the sizes of the matrices $\mathbf{U_s}$, $\mathbf{\Sigma}$, $\mathbf{F}$, $\mathbf{R}$, $\mathbf{Q}$ and $\hat{\mathbf{H}}$.

It is conceivable that an incorrect estimate of d will affect the success of the whole algorithm considerably. The channel matrix $\mathbf{H}$ must be of full column rank d. If it is close to being singular and hence ill-conditioned, the two singular value decompositions may not yield accurate results and, therefore, estimates of the signal subspace dimension d and the noise variance σ^2 may not be reliable. In such cases, the TXK algorithm more than likely will break down.

The SS algorithm, on the other hand, depends on the correct estimation of the signal subspace dimension $M + N$. The third step in the algorithm, which estimates the signal subspace dimension $M + N$, determines the sizes of the matrices $\mathbf{U_s}$, $\mathbf{U_n}$, and the filtering matrices $\mathbf{G}_i$, $M + N + 1 \leq i \leq s_r N$. Theoretically, an incorrect estimate of $M + N$ will affect the success of the whole algorithm. But the SS algorithm is different from the TXK algorithm in that it requires only one singular value decomposition while the TXK requires two. Hence a mismatch of the signal subspace dimension d will have more impact on the TXK algorithm than a mismatch of the signal subspace dimension $M + N$ affects the SS algorithm. Moreover, according to the authors of [19], it is even possible to recover the channel coefficients using only a few noise subspace eigenvectors. That is, the SS algorithm may turn out to be robust to mismatch of the signal subspace dimension $M + N$. However, this claim was not proved in full generality in [19]. This aspect of the algorithm requires further research.

In short, the accurate estimation of the signal subspace dimension (d for TXK or $M+N$ for the SS algorithm) seems to be important in the implementation of the TXK and SS algorithms.

5.0 COMPUTER SIMULATIONS

The impulse response $h(\cdot)$ of the multipath channel in our simulations is of the

form

$$h(t) = \sum_{j=1}^{P} c_j p(t - \tau_j) \tag{99}$$

where $p(\cdot)$ is the pulse-shape transmitted, P is the number of propagation paths, and τ_j and c_j are the time-delay and amplitude of the j-th path for $1 \leq j \leq P$. A shifted and truncated raised-cosine or square-root raised cosine pulse shape was used in tests. Specifically, the basic raised-cosine and square-root raised cosine pulse shapes are defined by

$$q_1(t, T_s, \beta) = \text{sinc}(t/T_s) \frac{\cos(\pi\beta t/T_s)}{1 - 4\beta^2 t^2/T_s^2} \tag{100}$$

and

$$q_2(t, T_s, \beta) = \frac{4\beta}{\pi\sqrt{T_s}} \frac{\cos((1+\beta)\pi t/T_s) + [\sin((1-\beta)\pi t/T_s)/(4\beta t/T_s)]}{1 - (4\beta t/T_s)^2}, \tag{101}$$

respectively, where T_s is the symbol period and β is the roll-off factor. Examples of these functions are illustrated in Fig. 1 for $\beta = 0.25$. The pulse shape used was either $p_1(t) = q_1^*(t - (L_h/2)T_s)$ or $p_2(t) = q_2(t - (L_h/2)T_s)$ where L_h is the duration of the channel impulse response in symbol periods, i.e., the impulse response is a finite causal function. The sampling period Δ was normalized to be 1. The modulation types were either 8-PSK or 16-QAM .

The simulations were divided into two groups. In the first group, a raised-cosine pulse shaping filter with roll-off factor $\beta = 0.25$ was used. Referring to Eq. 99, the number of paths was $P = 3$, with delays $\tau_1 = 0$, $\tau_2 = 2.5T_s$, $\tau_3 = 4T_s$ and weights of the paths were, respectively, $c_1 = 1$, $c_2 = 0.3$, $c_3 = -0.15$. The duration of the impulse response $h(\cdot)$ was set to $L_h = 6T_s$. The channel impulse response [17] $h(\cdot)$ used in the simulations can therefore be written as

$$h_1(t) = (c_1 p_1(t) + c_2 p_1(t - 2.5T_s) + c_3 p_1(t - 4T_s))W_{6T_s}(t) \tag{102}$$

where W_{6T_s} is the truncation window function defined by $W_{6T_s}(t) = 1$ if $0 \leq t \leq 6T_s$ and $W_{6T_s}(t) = 0$ if $t \leq 0$ or $6T_s \leq t$. The channel impulse response is illustrated in Fig. 2(a). Different oversampling ratios s_r were used, ranging from 8 to 32, depending on the experiment. Generally, it was found that higher ratios worked better. The length L

of the observation window for the TXK algorithm was $2T_s$ ($N = 2$) and the SS algorithm was $3T_s$ ($N = 3$). The the number of samples in the observation windows used therefore ranged from $s_r N = 24$ to $s_r N = 84$. The number N_1 of observation windows used in the estimation of $\mathbf{R}_x(0)$ and $\mathbf{R}_x(1)$ in the TXK algorithm (or $\mathbf{R}_x$ in the SS algorithm) was either 300 for 16-QAM modulations, or 100 for 8-PSK modulations. The signal to noise ratio SNR, defined as

$$\text{SNR} = 10\log_{10}\left(\frac{||r||^2}{||w||^2}\right) \tag{103}$$

where r and w are the signal and noise, respectively, and $||\cdot||^2$ refers to the mean-squared value, was set to a value of 30 dB in all tests. In the second group of simulations, a square-root raised-cosine pulse-shaping filter with roll-off factor $\beta = 0.25$ was used. All the other parameters were exactly the same as in the first group of simulations.

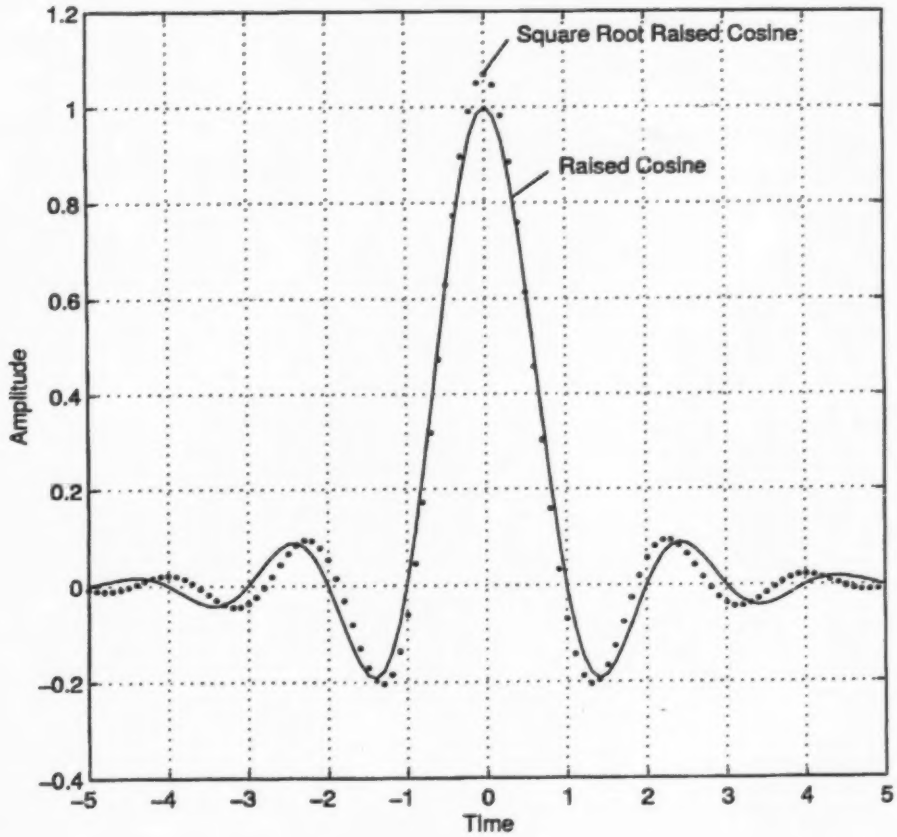


Figure 1: Examples of raised cosine and square-root raised cosine pulses for $\beta = 0.25$.

Symbols from 16-QAM and 8-PSK modulations were transmitted over the multipath channel defined above. Examples of the kinds of signal constellations one would obtain if the received signal is unequalized at the receiver are illustrated in Figs. 3(a) and (b) for the two modulations, respectively. Obviously, it is impossible to determine the constellation structure. The objective of the equalizer algorithms is to correct for the distortions introduced by the channel's multipath.

The TXK and SS algorithms recover the transmitted symbols on a block by block basis. A sufficiently large block N_1 of channel output sample vectors $\mathbf{X}_k$ are assembled followed by estimation of the correlation matrices $\mathbf{R}_x(0)$ and $\mathbf{R}_x(1)$. The channel matrix $\mathbf{H}$ is estimated using two SVD decompositions for the TXK algorithm and one for the SS algorithm. The estimated channel matrix $\hat{\mathbf{H}}$ is then used to recover the transmitted symbols through a matrix pseudo inversion operation which is equivalent to equalizing the channel. This process repeats itself for succeeding blocks of received samples.

The TXK and SS algorithms require an estimate of the signal subspace dimension. The signal subspace dimension d is estimated using either the AIC or MDL criterion discussed earlier. Typical examples of the eigenvalues of the autocorrelation matrix of the channel output data, and the values produced by the AIC and MDL critera are illustrated in Fig. 2(b), (c) and (d), respectively, for a 16-QAM signal, $SNR = 30$ dB, $s_r = 32$, and $N = 2$. The signal subspace dimension estimated by AIC and MDL was $d = 7$ for this example.

Figures 4 to 13 present a number of examples of equalized signal constellations using the TXK, SS and CMA algorithms.

Figures 4(a) to 4(d) show the equalized constellations resulting from use of the TXK algorithm when the modulation was 16-QAM, the symbol block size was 300 symbols, a raised cosine pulse was used and the observation window was $N = 2$. One general comment is that the constellations are rotated somewhat. This rotation is due to the arbitrary phase shift which the algorithm introduces to the estimated channel impulse response as discussed earlier. Figures 4(a), (b) and (c) illustrate the effect of the signal subspace parameter d on the results for a fixed oversampling ratio $s_r = 32$. It is clear that the constellation associated with the optimum value ($d = 7$, in Fig. 4(b)) produced the best results. Selecting a value greater than this ($d = 8$, in Fig. 4(c)) yielded poor results. Figure 4(d) shows the results for an oversampling rate which is half that used in Figs. 4(a) to (c) ($s_r = 16$) and the optimum signal subspace parameter, $d = 7$. The constellation can be made out, but is not as good as the one in Fig. 4(b). The comparison

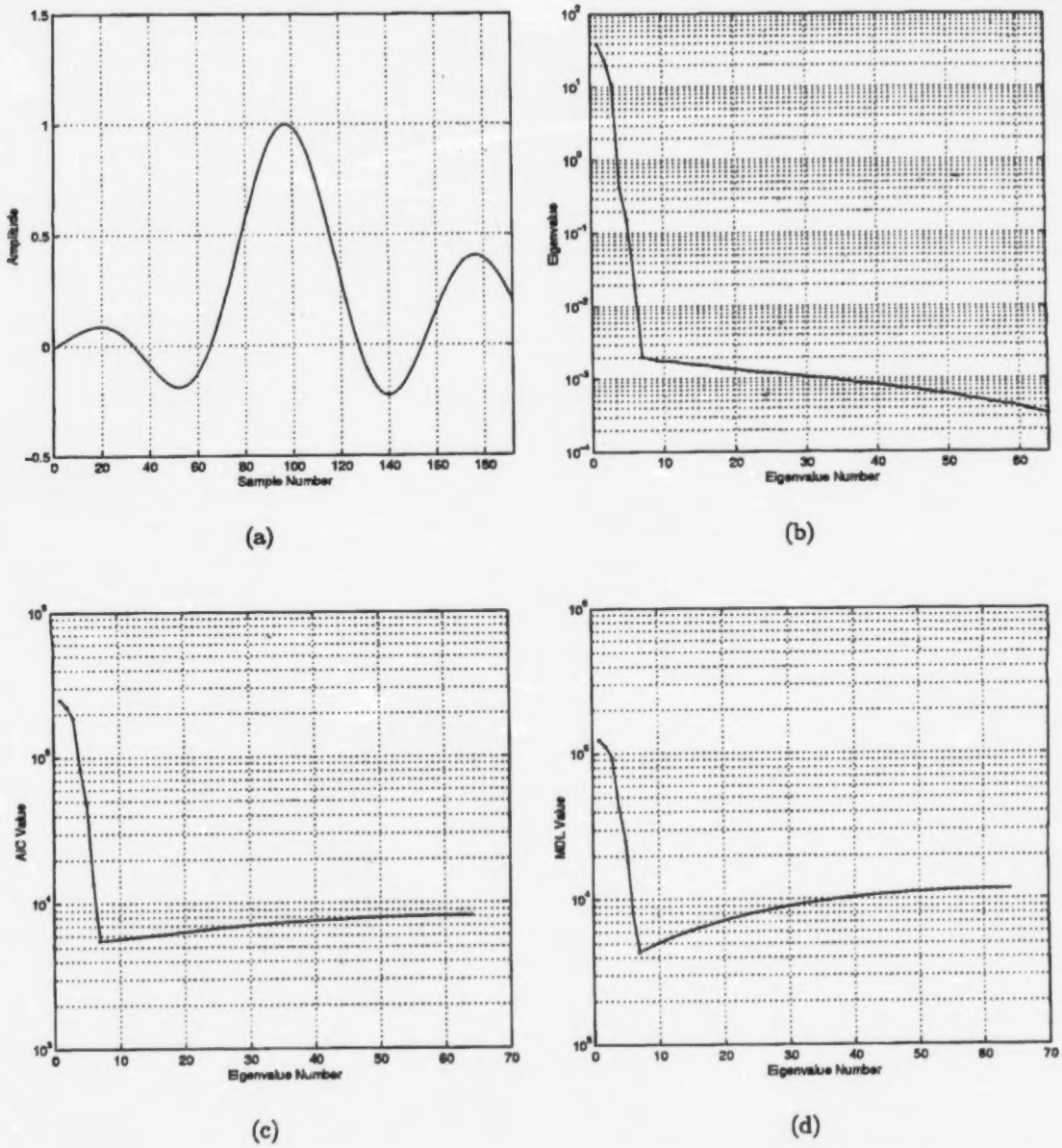
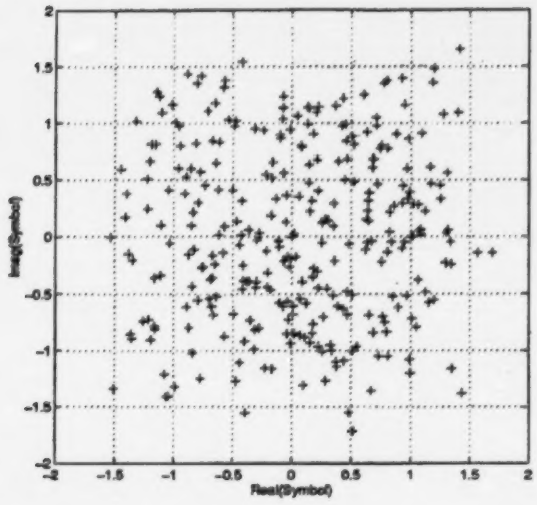
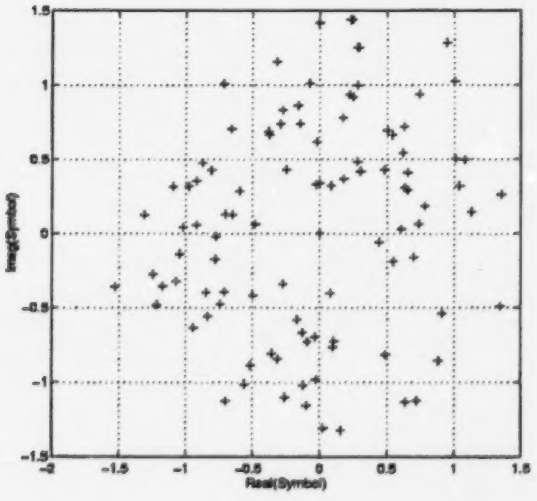


Figure 2: Three-path channel characteristics for $SNR = 30$ dB showing the (a) channel impulse response, (b) eigenvalues, (c) AIC values, and (d) MDL values.



(a)



(b)

Figure 3: Examples of signal constellations where the signal has not been equalized, $SNR = 30$dB (a) 16-QAM, (b) 8-PSK.

of Fig. 4(d) to Fig. 4(b) illustrates the performance enhancement which can be achieved by oversampling.

Figure 5 shows the constellations for the case when the pulse shape was of the square root raised cosine type. The same parameters as those used to produce the constellations in Fig. 4 were used. One can see that the results are not quite as good as those in the previous figure. It can only be surmised that, even though the square root raised cosine pulse is similar to the raised cosine one, the algorithm may have more difficulty in equalizing signals transmitted using this pulse shape.

Figures 6 and 7 illustrate the TXK equalized constellations for the case when the modulation was 8-PSK. Note however, that the block size used in these examples was 100 symbols compared to 300 symbols for the 16-QAM modulation. The algorithm obviously had less difficulty in equalizing this signal compared to the QAM signal. Figure 6(d) shows the results for an oversampling ratio 4 times less than the one used to produce the constellations in Figs. 6(a), (b), and (c). Even though the optimum signal subspace value of $d = 7$, the constellation is quite poor. However, when a value one less than the optimum value was used, the constellation was much better (not shown). Finally, as in the case of the QAM signal and the square root raised cosine pulse, the constellations for the 8-PSK signal were slightly noisier as shown in Fig. 7.

Figures 8 to 11 show the results for the SS algorithm. The parameters that were used to generate these results were slightly different than those used in the TXK algorithm. It was found that the latter's parameters (e.g., $N = 2$, $s_r = 32$) would not work as well for the SS algorithm. Nonetheless, the results for the SS algorithm were quite similar to those obtained from the TXK algorithm.

Finally, the constellations produced by the CMA algorithm are illustrated in Figs. 12 and 13 for the two modulation types. The main point to be made here is that the CMA algorithm took longer to converge to an identifiable constellation. For example, Fig. 12(b) shows the 16-QAM constellation starting to take shape after 1200 symbols, compared to 300 for the TXK and SS algorithms. Similarly, Fig. 13(b) shows the 8-PSK constellation starting to take shape after 600 symbols compared to 100 for the TXK and SS algorithms.

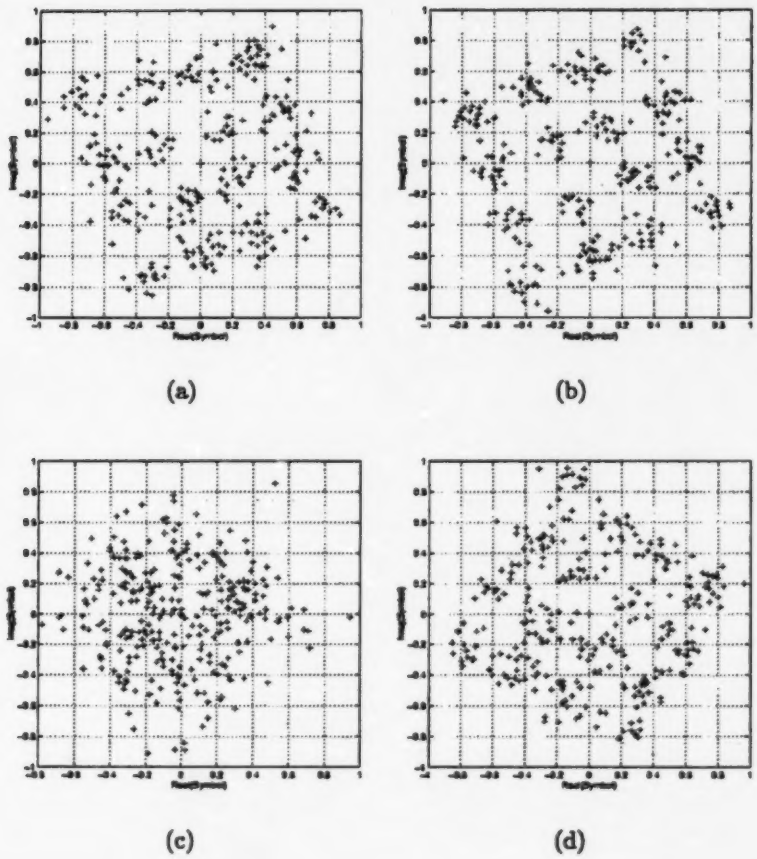


Figure 4: Equalized signal constellation (raised cosine pulse, 16-QAM modulation, block size 300 symbols, $SNR = 30$dB) using the TXK algorithm with the following parameters: (a) oversampling ratio $s_r = 32$, observation window size $N = 2$, signal subspace size $d = 6$; (b) $s_r = 32$, $N = 2$, $d = 7$ (optimum value for this set of parameters); (c) $s_r = 32$, $N = 2$, $d = 8$; (d) $s_r = 16$, $N = 2$, $d = 7$

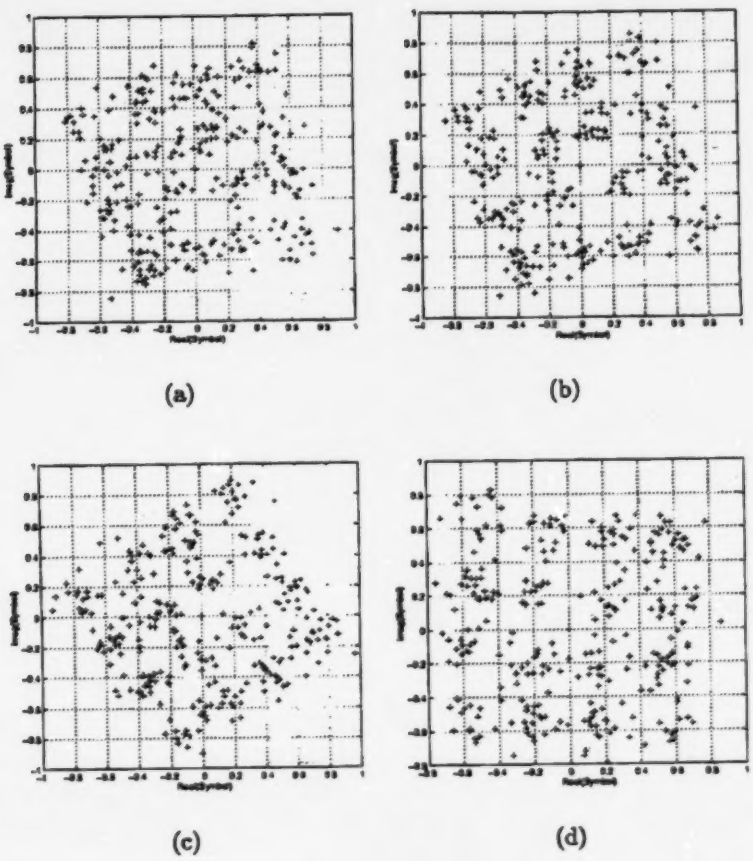


Figure 5: Equalized signal constellation (square-root raised cosine pulse, 16-QAM modulation, block size 300 symbols, $SNR = 30$dB) using the TXK algorithm with the following parameters: (a) oversampling ratio $s_r = 32$, observation window size $N = 2$, signal subspace size $d = 6$; (b) $s_r = 32$, $N = 2$, $d = 7$ (optimum value); (c) $s_r = 32$, $N = 2$, $d = 8$; (d) $s_r = 16$, $N = 2$, $d = 7$ (optimum value)

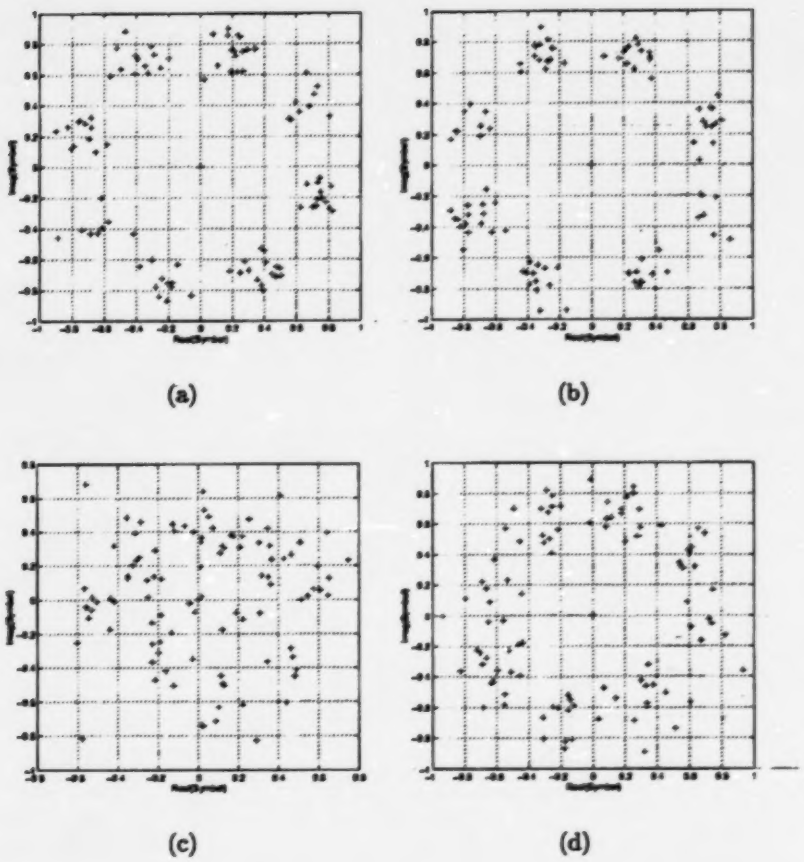


Figure 6: Equalized signal constellation (raised-cosine pulse, 8-PSK, 100 symbols, $SNR = 30$dB) using the TXK algorithm with the following parameters: (a) oversampling ratio $s_r = 32$, observation window size $N = 2$, signal subspace size $d = 5$; (b) $s_r = 32$, $N = 2$, $d = 6$ (optimum value); (c) $s_r = 32$, $N = 2$, $d = 7$; (d) $s_r = 8$, $N = 2$, $d = 7$ (optimum value)

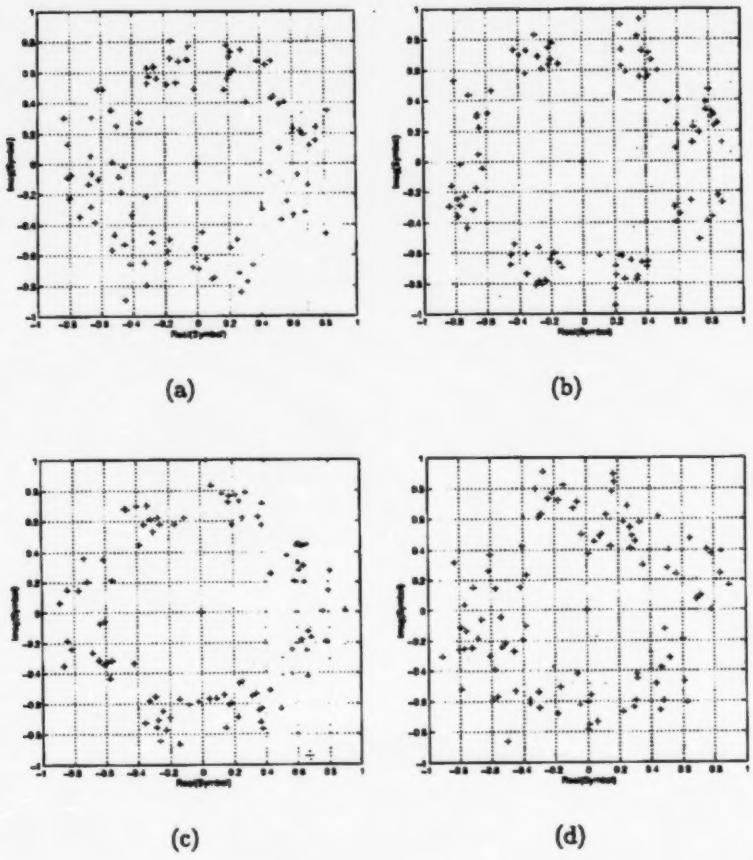


Figure 7: Equalized signal constellation (square-root raised cosine pulse, 8-PSK modulation, block size 100 symbols, $SNR = 30$dB) using the TXK algorithm with the following parameters: (a) oversampling ratio $s_r = 32$, observation window size $N = 2$, signal subspace size $d = 5$; (b) $s_r = 32$, $N = 2$, $d = 6$ (optimum value); (c) $s_r = 32$, $N = 2$, $d = 7$; (d) $s_r = 16$, $N = 2$, $d = 7$.

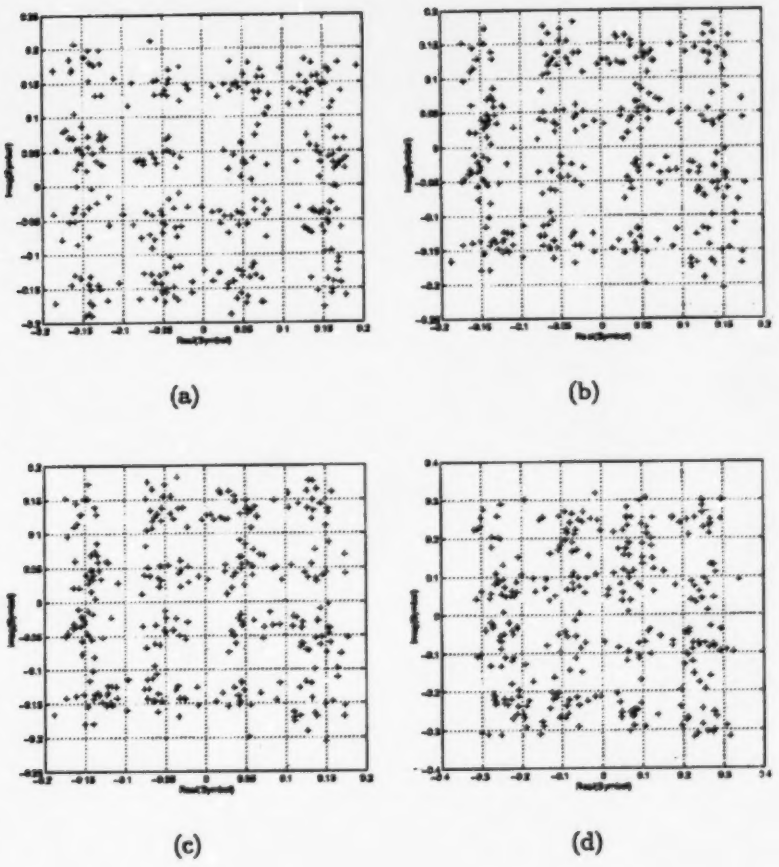


Figure 8: Equalized signal constellation (raised-cosine pulse, 16-QAM, block size 300 symbols, $SNR = 30$ dB) using the SS algorithm: (a) oversampling ratio $s_r = 28$, observation window size $N = 3$, signal subspace size, $d = 7$; (b) $s_r = 28$, $N = 3$, $d = 8$; (c) $s_r = 28$, $N = 3$, $d = 9$; (d) $s_r = 8$, $N = 3$, $d = 8$ (optimum value)

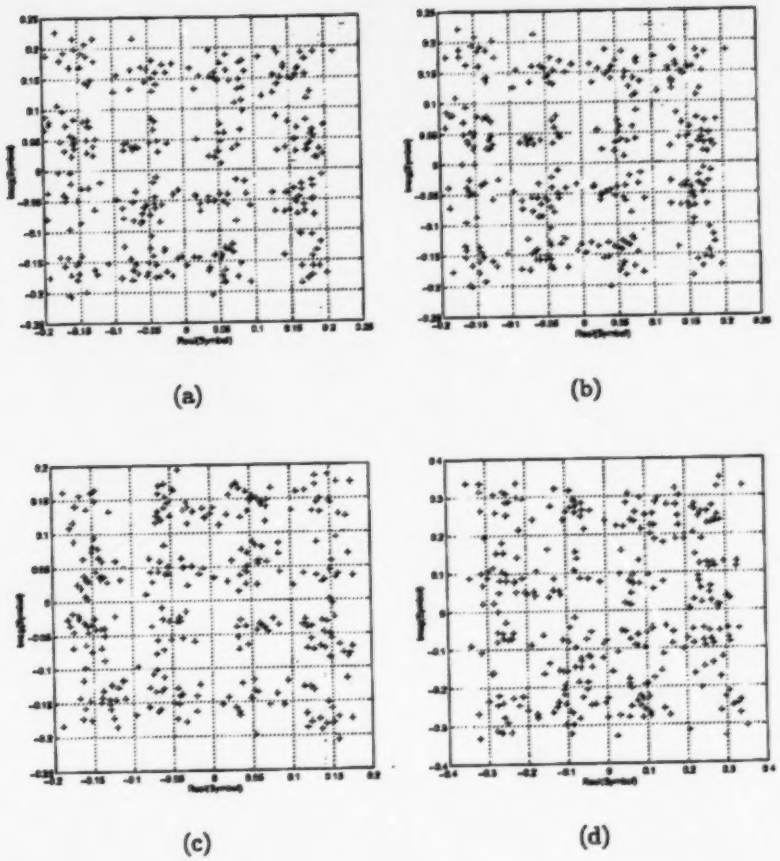


Figure 9: Equalized signal constellation (square-root raised cosine pulse, 16-QAM modulation, block size 300 symbols, $SNR = 30$dB) for SS algorithm with the following parameters: (a) oversampling ratio $s_r = 28$, observation window size $N = 3$, signal subspace size $d = 7$; (b) $s_r = 28$, $N = 3$, $d = 8$ (optimum value); (c) $s_r = 28$, $N = 3$, $d = 9$; (d) $s_r = 8$, $N = 3$, $d = 8$ (optimum value)

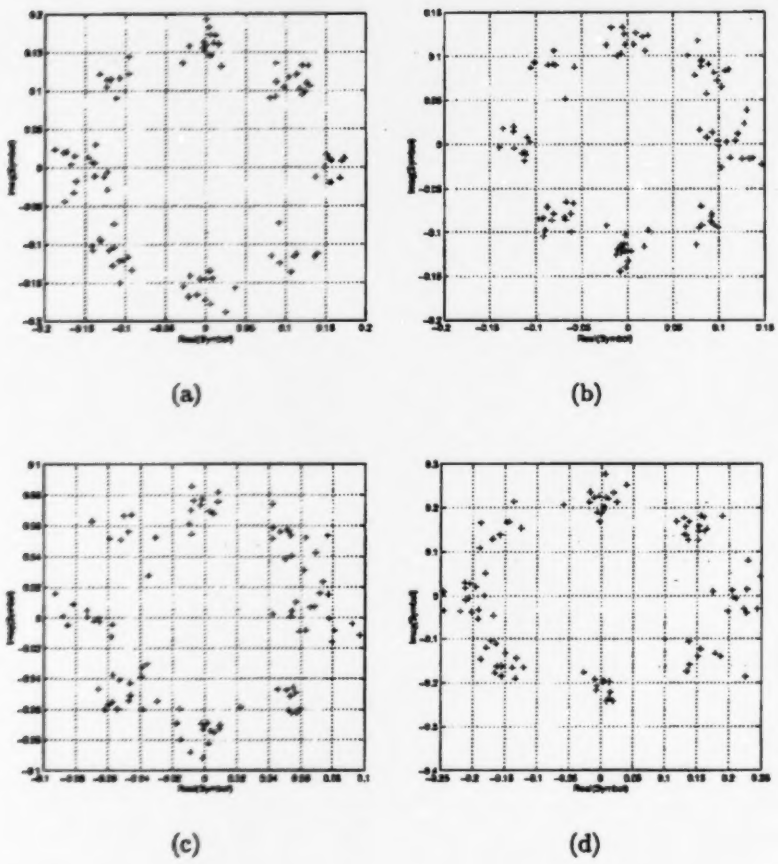


Figure 10: Equalized signal constellation (raised-cosine pulse, 8-PSK, block size 100 symbols, $SNR = 30$ dB) using the SS algorithm with the following parameters: (a) oversampling ratio $s_r = 28$, observation window size $N = 3$, signal subspace size, $d = 7$; (b) $s_r = 28$, $N = 3$, $d = 8$ (optimum value); (c) $s_r = 28$, $N = 3$, $d = 9$; (d) $s_r = 8$, $N = 3$, $d = 8$ (optimum value)

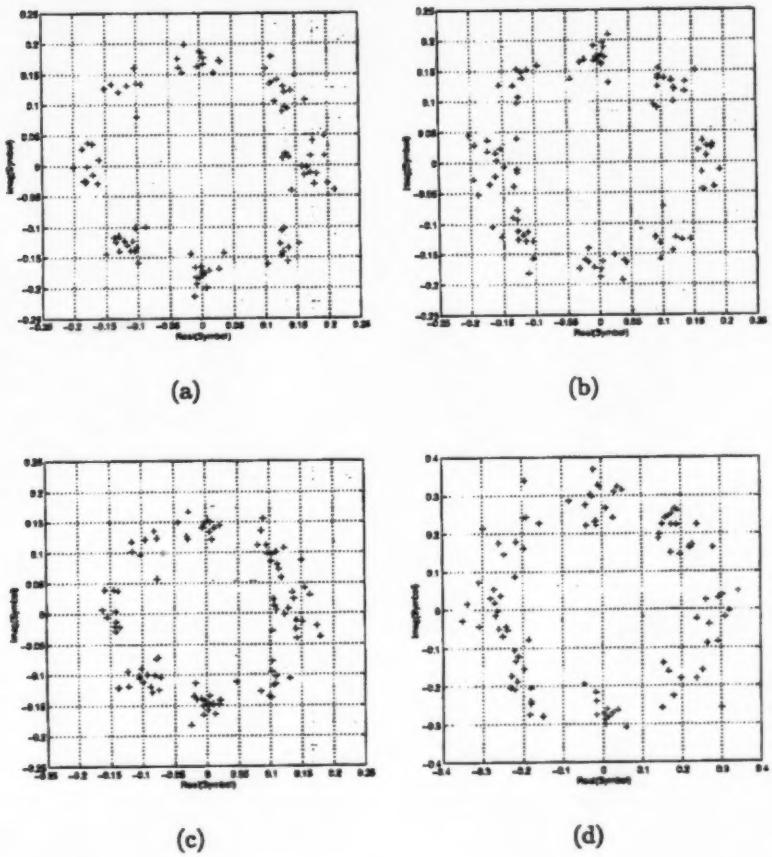


Figure 11: Equalized signal constellation (square-root raised cosine pulse, 8-PSK modulation, block size 100 symbols, $SNR = 30$dB) using the SS algorithm with the following parameters: (a) oversampling ratio $s_r = 28$, observation window size $N = 3$, signal subspace size $d = 7$; (b) $s_r = 28$, $N = 3$, $d = 8$ (optimum value); (c) $s_r = 28$, $N = 3$, $d = 9$; (d) $s_r = 8$, $N = 3$, $d = 9$ (optimum value)

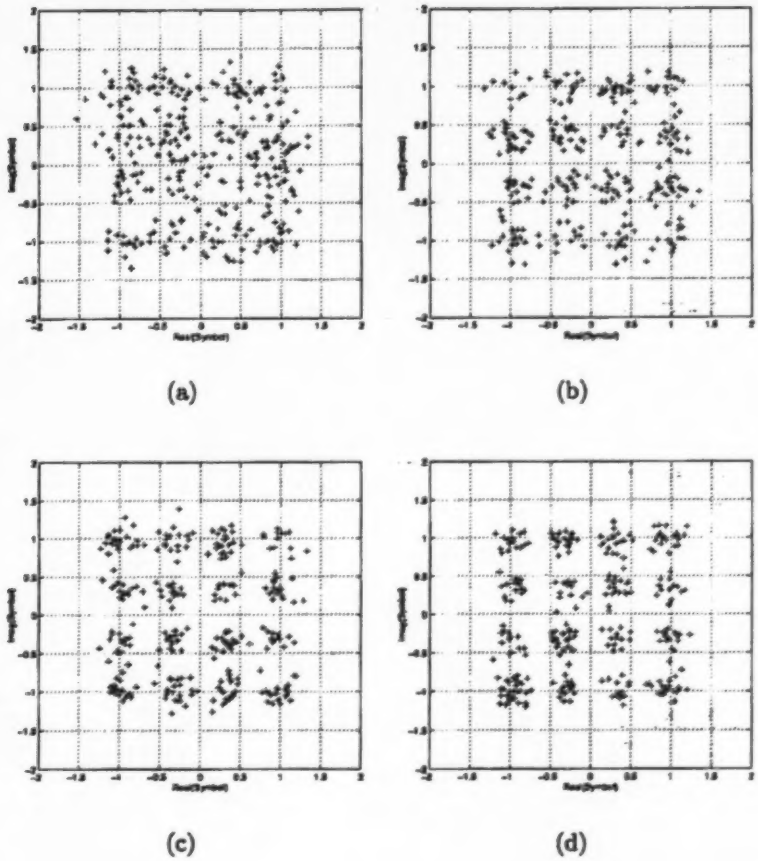


Figure 12: Equalized signal constellation (raised cosine pulse, equalizer length equal to 12 taps, 16-QAM modulation, $SNR = 30$dB, oversampling ratio $s_r = 4$) using the CMA algorithm, after: (a) 900 symbols, (b) 1200 symbols, (c) 1500 symbols, (d) 1800 symbols.

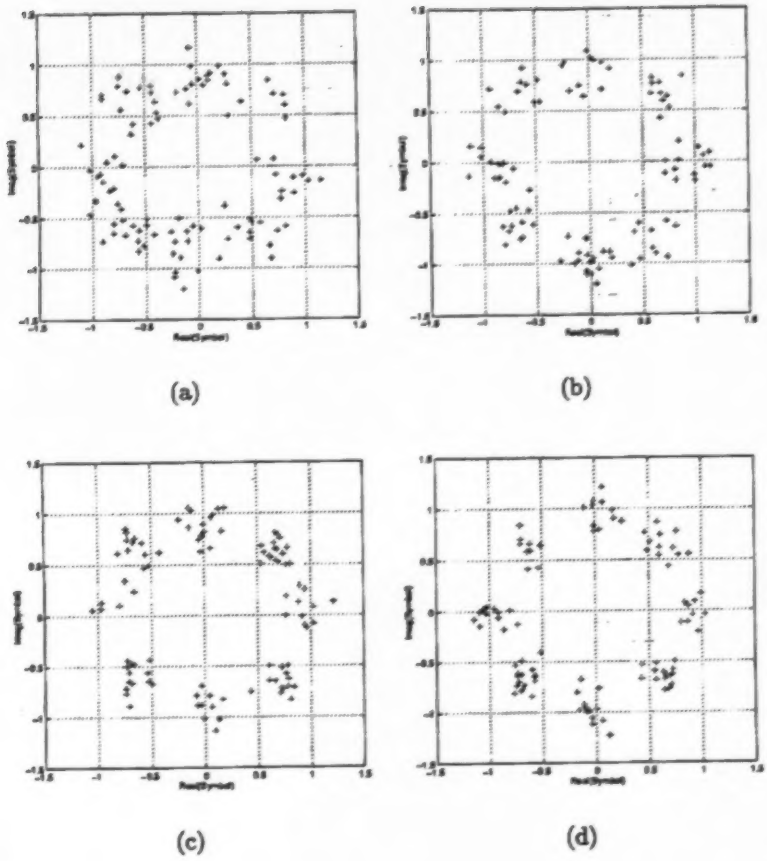


Figure 13: Equalized signal constellation (raised cosine pulse, equalizer length equal to 12 taps, 8-PSK modulation, $SNR = 30$dB, oversampling ratio $s_r = 4$) using the CMA algorithm, after: (a) 500 symbols, (b) 600 symbols, (c) 700 symbols, (d) 800 symbols.

6.0 CONCLUDING REMARKS

The results presented in this report demonstrate that the fast convergent equalizer algorithms can equalize channels faster than the CMA algorithm by up to a factor of four times. The main problem with them, however, is their lack of robustnes compared to the CMA algorithm, and the significant computational effort required to obtain any meaningful results. For example, high oversampling ratios were required to produce good equalization results. As far as the authors are aware, results using higher sampling ratios have not been published in the open literature. It was found that the performance of the algorithms can be sensitive to the estimated signal subspace parameter, particularly for low oversampling ratios. At times the algorithms will work well using the estimated value, other times a slightly lower value would produce superior results. The algorithms, in their current formulation at least, are not practical for online applications, but may be useful for offline signal processing tasks where fast convergence is required. An increase in convergence rate is significant, particularly when attempting to equalize a signal that has been transmitted over a (rapidly) time-varying channel.

REFERENCES

[1] Y. Sato, "A Method of Self-Recovering Equalization for Multi-level Amplitude Modulation", IEEE Transactions on Communications, Com-23, pp. 679-682, June 1975

[2] D. N. Godard, "Self-Recovering Equalization and Carrier Tracking in Two-Dimensional Communication Systems ", IEEE Transactions on Communications, Com-28, pp. 1867-1875, November 1980

[3] L. Tong and S. Perreau, "Multichannel Blind Identification: From Subspace to Maximum Likelihood Methods ", Proceedings of the IEEE, vol. 86, October, 1998

[4] L. Tong, G. Xu and T. Kailath, "A New Approach to Blind Identification and Equalization of Multipath Channels", in Proc. 25th Asilomar Conf, pp. 856-860

[5] L. Tong, G. Xu and T. Kailath, "Blind Identification and Equalization Based on Second-Order Statistics: A Time Domain Approach", IEEE Transactions on Information Theory, vol. 40, March 1994, pp. 340-349

[6] H. Liu, G. Xu, L. Tong and T. Kailath, "Recent Developments in Blind Channel Equalization: From Cyclostationarity to Subspaces", Signal Processing, vol. 50, pp. 83-99, 1996

[7] S. Haykin (Ed.), "Blind Deconvolution", Englewood Cliffs, New Jersey, Prentice-Hall, 1994

[8] D. L. Donoho, "On Minimum Entropy Deconvolution", Applied Time Series Analysis II, D. F. Findley, Ed., New York, Academic Press, 1981

[9] J. R. Treichler and B. G. Agee, "A New Approach to Multipath Correction of Constant Modulus Signals", IEEE Transactions on Acoustics, Speech and Signal Processing, vol. ASSP-31, April, 1983

[10] O. Shalvi and E. Weinstein, "New Criteria for Blind Deconvolution of Nonminimum Phase Systems (Channels) ", IEEE Transactions on Information Theory, vol. 36, pp. 312-321, March 1990

[11] O. Shalvi and E. Weinstein, "Super-exponential Methods for Blind Deconvolution", IEEE Transactions on Information Theory, vol. 39, March 1993

[12] O. Shalvi and E. Weinstein,"Universal Methods for Blind Deconvolution", in Blind Deconvolution, S. Haykin, Ed., Englewood Cliffs, New Jersey, Prentice-Hall, 1994

[13] D. Hatzinakos and C. Nikias, "Estimation of Multipath Channel Response in Frequency Selective Channels", IEEE Journal on Selected Areas of Communications, vol. SAC-7, pp. 12-19, January, 1989

[14] D. Hatzinakos and C. Nikias, "Blind Equalization Using a Triceptrum-based Algorithm", IEEE Transactions on Communications, May, 1991

[15] G. Giannakis and J. Mendel, "Identification of Nonminimum Phase Systems Using Higher Order Statistics", IEEE Transactions on Acoustics, Speech, Signal Processing, vol. ASSP-37, pp. 360-377, March, 1989

[16] C. R. Johnson, Jr., P. Schniter, T. J. Endres, J. D. Behm, D. R. Brown, and R. A. Casas, "Blind Equalization Using the Constant Modulus Criterion: A Review ", Proceedings of the IEEE, October, 1998

[17] B. W. Kozminchuk and X. Huang, "Joint Blind Synchronization of M-PSK and M-QAM Signals", DREO Technical Note 96-004, 1996

[18] B. W. Kozminchuk and Z. Hu, "Blind Equalization and Synchronization of M-PSK and M-QAM Signals in a Stationary Multipath Environment", DREO Technical Note 97-011, 1997

[19] E. Moulines, P. Duhamel, J-F. Cardoso and S. Mayrargue, "Subspace Methods for the Blind Identification of Multichannel FIR Filters", IEEE Transactions on Information Theory, vol. 40, March 1994, pp. 340-349

[20] M. Wax and T. Kailath, "Detection of Signals by Information Theoretic Criterion", IEEE Transactions on Acoustics, Speech, Signal Processing, vol. ASSP-33, No. 2, pp. 387-392, April, 1985